

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2
Perfect score: 2239
Sequence: 1 MSVGCPEPPRSLTCCGPG.....GRADSKGQVVLATAICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp.*
- 5: /cgn2_6/ptodata/1/iaa/ECTUS.COMB.pdp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	4	US-09-816-094-2
2	2084.5	93.1	417	4	US-09-816-094-4
3	341.5	15.3	729	2	US-08-677-298-2
4	341.5	15.3	729	4	US-09-523-849-33
5	340	15.2	513	3	US-08-924-183-8
6	340	15.2	513	4	US-09-488-364-8
7	336	15.0	765	4	US-09-975-326-4
8	335.5	15.0	766	4	US-09-975-326-2
9	334	14.9	793	4	US-09-523-849-32
10	333.5	14.9	779	4	US-08-817-832B-31
11	333	14.9	476	3	US-08-924-183-1
12	333	14.9	476	4	US-09-488-364-2
13	328.5	14.7	448	2	US-09-159-385-2
14	328.5	14.7	448	3	US-09-186-277-2
15	328	14.6	476	3	US-09-522-800-16
16	328	14.6	476	3	US-08-924-183-1
17	328	14.6	476	4	US-09-488-364-1
18	328	14.6	476	4	US-09-113-785-1
19	323.5	14.4	722	4	US-09-984-890-4
20	323.5	14.4	722	4	US-08-817-832B-32
21	322.5	14.4	483	3	US-08-924-183-9
22	322.5	14.4	483	4	US-09-488-364-9
23	322.5	14.4	745	4	US-09-523-849-36
24	321.5	14.4	724	4	US-09-984-890-2
25	318	14.2	454	2	US-09-159-385-1
26	318	14.2	454	3	US-09-186-277-1
27	315	14.1	552	3	US-08-557-006C-40

28	311.5	13.9	582	4	US-09-428-711A-2	Sequence 2, Appli
29	311.5	13.9	1051	4	US-09-428-711A-14	Sequence 14, Appli
30	310	13.8	658	4	US-09-930-181-2	Sequence 2, Appli
31	309	13.8	633	3	US-08-557-006C-43	Sequence 43, Appli
32	308.5	13.8	257	2	US-07-857-224B-25	Sequence 25, Appli
33	307	13.7	543	4	US-09-529-093A-2	Sequence 2, Appli
34	307	13.7	543	4	US-09-529-154-2	Sequence 2, Appli
35	305.5	13.6	476	3	US-08-924-183-7	Sequence 7, Appli
36	305.5	13.6	476	4	US-09-488-364-7	Sequence 7, Appli
37	302	13.5	631	4	US-09-579-664B-11	Sequence 11, Appli
38	302	13.5	776	4	US-09-523-849-34	Sequence 34, Appli
39	298.5	13.3	345	3	US-09-101-146-1	Sequence 1, Appli
40	298	13.3	604	4	US-09-523-849-35	Sequence 35, Appli
41	295.5	13.2	257	3	US-09-101-146-6	Sequence 6, Appli
42	294.5	13.2	396	4	US-09-801-876B-2	Sequence 2, Appli
43	294.5	13.2	1050	4	US-09-428-711A-16	Sequence 16, Appli
44	294	13.1	436	4	US-09-734-673-2	Sequence 2, Appli
45	294	13.1	436	4	US-09-523-849-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-816-094-2

; Sequence 2, Application US/09816094

; Patent No. 6534299

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: THEREOF

; FILE REFERENCE: CLO00536-CIP

; CURRENT APPLICATION NUMBER: US/09/816,094

; CURRENT FILING DATE: 2001-03-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Human

; US-09-816-094-2

Query Match

Best Local Similarity 100.0%; Score 2239; DB 4; Length 424;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMOALTTLTAAADVTKHVELVRELG	60
Db	1	MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMOALTTLTAAADVTKHVELVRELG	60
Qy	61	KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFIKVDVVFET	120
Db	61	KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFIKVDVVFET	120
Qy	121	EDCVFAQYAPAGDLFDIIPPQVGLPEDTVKRCVQQLGLALDFMHRQVLRDIKENV	180
Db	121	EDCVFAQYAPAGDLFDIIPPQVGLPEDTVKRCVQQLGLALDFMHRQVLRDIKENV	180
Qy	181	LLFDRECRVKLADFGMTRRVGRVKVSGTIPTTAPEVCOAGRADGLAVDTGVDVWAFG	240
Db	181	LLFDRECRVKLADFGMTRRVGRVKVSGTIPTTAPEVCOAGRADGLAVDTGVDVWAFG	240
Qy	241	VLIFCVLTGNFPEAASGADAFEEFVRWGRGLPGLPSQWRPFTEPALRMFORLLALEP	300
Db	241	VLIFCVLTGNFPEAASGADAFEEFVRWGRGLPGLPSQWRPFTEPALRMFORLLALEP	300
Qy	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLEAPGLKRTVLTESG	360
Db	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLEAPGLKRTVLTESG	360
Qy	361	SGSRPAPPAGVSVPLVPVVPVVPVPEPGLAPGPPGRTDGRADSKGQVVLATAI	420

5/17/02
60205228

RESULT 4
US-09-523-849-33
; Sequence 33, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 Seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPEPRSLTCCGPG.....GRADKSGQVLTATAICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	424	4	US-09-816-094-2
2	2084.5	93.1	417	4	US-09-816-094-4
3	341.5	15.3	729	2	US-08-677-298-2
4	341.5	15.3	729	4	US-09-523-849-33
5	340	15.2	513	3	US-08-924-183-8
6	340	15.2	513	4	US-09-488-364-8
7	336	15.0	765	4	US-09-975-326-4
8	335.5	15.0	765	4	US-09-975-326-2
9	334	14.9	793	4	US-09-523-849-32
10	333.5	14.9	779	4	US-08-817-832B-31
11	333	14.9	476	3	US-08-924-183-2
12	333	14.9	476	4	US-09-488-364-2
13	328.5	14.7	448	2	US-09-159-385-2
14	328.5	14.7	448	3	US-09-186-277-2
15	328	14.6	476	3	US-09-522-800-16
16	328	14.6	476	3	US-08-924-183-1
17	328	14.6	476	4	US-09-488-364-1
18	328	14.6	476	4	US-09-113-785-1
19	323.5	14.4	722	4	US-09-984-890-4
20	323.5	14.4	722	4	US-08-817-832B-32
21	322.5	14.4	483	3	US-08-924-183-9
22	322.5	14.4	483	4	US-09-488-364-9
23	322.5	14.4	745	4	US-09-523-849-36
24	321.5	14.4	724	4	US-09-984-890-2
25	318	14.2	454	2	US-09-159-385-1
26	318	14.2	454	3	US-09-186-277-1
27	315	14.1	552	3	US-08-557-006C-40

ALIGNMENTS

RESULT 1

US-09-816-094-2

; Sequence 2, Application US/09816094

; Patent No. 6534299

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: THEREOF

; FILE REFERENCE: CL000536-CIP

; CURRENT FILING DATE: 2001-03-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Human

US-09-816-094-2

Query Match	Best Local Similarity	Score	DB 4;	Length	424;
Matches	424;	Conservative	0;	Mismatches	0;
Gaps	0;	Indels	0;	Gaps	0;
Qy	1	MSVGCPEPEPRSLTCCGPGTAPGPGAGVPLLTEDMOALTIRTLAASDVTKHVELREL	60		
Db	1	MSVGCPEPEPRSLTCCGPGTAPGPGAGVPLLTEDMOALTIRTLAASDVTKHVELREL	60		
Qy	61	KGTGKVDLVVYKGTGTMKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET	120		
Db	61	KGTGKVDLVVYKGTGTMKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET	120		
Qy	121	EDCVFAQAEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKENV	180		
Db	121	EDCVFAQAEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKENV	180		
Qy	181	LLFDRECRVKLADFGMTRVRCVGVKVSCTIPTYTAPEVCOAGRADGLAVDTGVVWAFG	240		
Db	181	LLFDRECRVKLADFGMTRVRCVGVKVSCTIPTYTAPEVCOAGRADGLAVDTGVVWAFG	240		
Qy	241	VLIFCVLTGNFPWEAASGADAFEEFVRWQGRGLPLPSQWRFFTEPALMFORLLALEP	300		
Db	241	VLIFCVLTGNFPWEAASGADAFEEFVRWQGRGLPLPSQWRFFTEPALMFORLLALEP	300		
Qy	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLPAGPLKRTVLTESG	360		
Db	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLPAGPLKRTVLTESG	360		
Qy	361	SGSRPAPPAVGSPLPVFPVVPVPEPGIACQGPRTDGRADKSGQVLTATAICV	420		

5/17/09
60205228

Db 60 IVKLYD-VFEDDDKLYVMEYCYGSGDLFDLKKRGLSEDEARFYARQILSALEYLHSQG 118
QY 170 LVHRDKPENVLFFDRECRVVKLADFGMTFRV--GCRVKRVSGTTPYTAPEVCQAGRAD 226
Db 119 IHRDLKPENILL-DSG-GHVKLADGLAKQLDSGGTLLTTFVTGYVAPEVL-IGKGY 175
QY 227 GLAVDTGVDVWAFGVLFVCLTGNFPWEAASGADAFPEEFVVRWQGRGLPGLPSQWRRT 286
Db 176 GKX---VDIWSLGVLYELLTGKPPF---PGDDQLLALFKKIGKPPPPPPPEW-KISP 227
QY 287 PALRMFORLLALEPERRGPAKE 308
Db 228 EAXDLKKLLVXDPKRLTAEE 249

RESULT 15
US-09-976-782-30
; Sequence 30, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/252,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Consensus
; OTHER INFORMATION: sequence
US-09-976-782-30

Query Match 15.6%; Score 350; DB 12; Length 254;
Best Local Similarity 35.9%; Pred. No. 7.4e-18;
Matches 97; Conservative 44; Mismatches 105; Indels 24; Gaps 10;
QY 53 YELVRELGKGYGVKVDLVVYKGTGKALKFNKKS--KTKLKNELREVSITNSLSSSPFI 110
Db 1 YELGKLGSGAFGVKYGKHKHDTGEIVAKTKIKRKSJSEKKRFLREIQILRL-SHPNI 59
QY 111 IKVFDVWPETEDCVFAQVAPAGLDLFIIPQ-VGLPEDTVKRCVQQLGLALDFMEGRQ 169
Db 60 VRLLG-VFEEDHLYVMEYEGGDLFDYLRNGLLLSEKAKKIALQILRGLVYLSRG 118
QY 170 LVHRDKPENVLFFDRECRVVKLADFGMTFRVGC-----RVKRVSGTTPYTAPEVCQAGRA 225
Db 119 IVHRDLKPENILL--DENGTVKIADFGGLARKLESSSYEKLATTFVGTPEYMAPEV-----L 171

QY 226 DGLAVDTGVDVWAFGVLFVCLTGNFPWEAASGADAFPEEFVVRWQGRGLPGLPSQWRRT 285
Db 172 EGRGYSSKVDVMSLGVLYELLTGKLPFPFGIDPLELFRKER-PRRLPLPPN-----CS 226
QY 286 EPALRMFORLLALEPERRGPAKEVREFLKH 315
Db 227 BELKDLIKKLNKDPKRPRTAKEI---LNH 253

Search completed: December 5, 2003, 10:05:22
Job time : 39 secs

Best Local Similarity 37.5%; Pred. No. 4.6e-21;	
Matches 101; Conservative 50; Mismatches 96; Indels 22; Gaps 11;	
Qy	53 YELVRELCKGYGKGVLDLVVYKGTGTMALKFVNKSKTKLK--NFLREVSTINSLSSPPF 109
Db	1 YELLEVLGKGFAGKGVYLARDKDTGRLVAKVIVKKEUKKKKKRERILREIKLKKL-DHPN 59
Qy	110 LKVFDDVVFFETDCVFAQEYAPAGDIFDIIPQVGLPEDTVKRCVQQLGALDFMHGRQ 169
Db	60 IVKLYD-VFEDDDKLYLWMEYCEGGDLFDLLKKRGLSEEARFVARQILSALEVLHSQG 118
Qy	170 LVHEDIKPENULLFDRECRRVKLADFGMTRVY---GRVKVSGTIPYTADEVCAQRAD 226
Db	119 IIRHDLKPENILL-DSD-GHVKLADFLGAKQLDSGGTLLTFVGTPEYMAPEVL-LQKGY 175
Qy	227 GLAVDTGVDVMAFGVLIPCVLTGNFFPMEASGADAFEEFVRWGRGLPGLPSONRRPTE 286
Db	176 GKA---VDINSLGVIIYELLITGRPPP---FGDQLLALFKKIGKPKPPFPPEPW-KLSP 227
Qy	287 PALRMFORLLALEPERRGPAKEVFRFTLKH 315
Db	228 EAKDLIKKLAVKQDEKSLTAEPA---LKH 253

```

RESULT 13
US-09-823-187-96
; Sequence 96, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Murallidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-96

Query Match          17.4%;   Score 390;   DB 11;   Length 256;
Best Local Similarity 36.8%;   Pred. NO. 9e-21;

```

[illegible]

```

RESULT 14
US-09-976-782-31
; Sequence 31, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Consensus
; OTHER INFORMATION: sequence
US-09-976-782-31

```

```

Query Match      17.1%; Score 383.5; DB 12; Length 250;
Best Local Similarity 37.4%; Pred.No.2.6e-20;
Matches          98; Conservative 49; Mismatches 96; Indels 19; Gaps 10;

QY 53 YELVRILGKGYGKVDLVVYKGTGKMGALFVNKSKTKLK--NFLREVSTINSLSSSPF 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 YELLEVILGKGFQGVKVLARDKTKGKLVAKIVIKKEKIKKKRRERILKEITKEL--DHPN 59

QY 110 IIKVDFVVFETECYVFAQRYAPAGLDLFIIPQVGLPEDIKTRCVQOGLGALDFMGRO 169

```

```
RESULT 11
US-10-353-690-50
; Sequence 50, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodrigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MP102-018P1NOMNIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-353-690-50

Query Match 24.4%; Score 546.5; DB 12; Length 348;
Best Local Similarity 42.9%; Pred. No. 5e-32;
Matches 117; Conservative 42; Mismatches 111; Indels 3; Gaps 3;

QY 34 EDMQALTURTAAASDVTKHYELVRELKGTGKVDLVVYKGTWKALKFVKNKSKTKLN 93
DB 43 EDMWTLSQTLVRAEVDLEYEVRPLGQRYGRVLLVTHRQKGTPLALKQPKPRTSRG 102
QY 94 FLREVSITNLSSSPFIKFDVWFETEDCVYFAQYAPAGDLFIIPQVGLPEDTVYKR 153
DB 103 FLYEFCVGLSLGAHSAIVTAYGIGIESAHSYFLTEPVLHGLDMAFIQPKVGLPQPAVHR 162
QY 154 CVQQGLGLDLMHGRQLVHRDIKPNVLLFDECRVRKLAIFGMTRVRGCRVKSRTIP 213
DB 163 CAAQJASALEYTHARGLVRLDKPENVLVCPACRRFKLTDFTGTRPRGTLLRLAGPPIP 222
QY 214 YTAPEVC-QAGRADGLAVDTGVDVWAFGLVFCVLITGNFPWE-AASGADAFEEFVRQR 271
```

```
Db 223 YTAPELCAPPPLPEGLPIQPALDAWALGVLLFCLLTGYFFPDRPLAEDAFYEDFLIWA 282
QY 272 GRUP-GLPSOWRRRTTEPALRMFORLLALEPERR 303
Db 283 SGQRDRPQFWFGLAPADALLRGLLDHPERR 315

RESULT 12
US-09-863-776-40
; Sequence 40, Application US/09863776
; Publication No. US20030198953A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spaderna, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Li, Li
; APPLICANT: Taupier, Raymond J
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: No. US20030198953A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-020
; CURRENT APPLICATION NUMBER: US/09/863,776
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/206,679
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,688
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,829
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/207,748
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/207,798
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/208,263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/208,831
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/209,451
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 60/210,060
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/219,507
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/221,337
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/221,927
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/263,135
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,688
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,694
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Serine/threonine protein kinase, catalytic domain
; OTHER INFORMATION: Consensus Sequence
US-09-863-776-40

Query Match 17.6%; Score 394; DB 12; Length 256;
```

Db 181 LLEDFECRRVXKLDGFMTRVRCRVKRVSGTIPYTAPEVCQAGRADGFAVDTGVVDVWAFG 240
QY 241 VLIFFCVLTGNFPWEAASGADAFFEFVWQGRPLFGLPSQWRRFTEPALRMFORLLALEP 300
Db 241 VLIFFCVLTGNFPWEAASGADAFFEFVWQGRPLFGLPSQWRRFTEPALRMFORLLALEP 300
QY 301 ERGGPAKEVVFRLKHELTSELRRRSHRARKPPGDRPPAAGPLRLLEAGPLKRTVLTESG 360
Db 301 ERGGPAKEVVFRLKHELTSELRRRSHRARKPPGDRPLP--GPLRLLEAGPLKRTVLTESG 358
QY 361 SGSRPAPVAGSVPLFVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
Db 359 SGSRPSPSVGPFV-----VPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 413
QY 421 EICV 424
Db 414 EICV 417

RESULT 9
US-09-916-790-2
; Sequence 2, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Sifos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916, 790
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-916-790-2

Query Match 81.4%; Score 1822; DB 9; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.7e-125;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 MALFVNKSKTKLKNFLREVSTNSLSPPFIKVFVDFVVFETEDCYVFAQYAPAGDLFD 138
Db 1 MALFVNKSKTKLKNFLREVSTNSLSPPFIKVFVDFVVFETEDCYVFAQYAPAGDLFD 60
QY 139 IIPQVGLPDTVKRCVQOGLGLDFHGRQLVHRDIKPEVNLVDFRECRVKLADFGMT 198
Db 61 IIPQVGLPDTVKRCVQOGLGLDFHGRQLVHRDIKPEVNLVDFRECRVKLADFGMT 120
QY 199 RVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVVDVWAFGLVLCVLTGNFPWEAASG 258
Db 121 RVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVVDVWAFGLVLCVLTGNFPWEAASG 180
QY 259 ADAFFEFVWQGRPLFGLPSQWRRFTEPALRMFORLLALEPERRGPAKEVFRFLKHELT 318
Db 181 ADAFFEFVWQGRPLFGLPSQWRRFTEPALRMFORLLALEPERRGPAKEVFRFLKHELT 240
QY 319 SELRRRSHRARKPPGDRPPAAGPLRLLEAGPLKRTVLTESGSRPAPVAGSVPLFVP 378
Db 241 SELRRRSHRARKPPGDRPPAAGPLRLLEAGPLKRTVLTESGSRPAPVAGSVPLFVP 300
QY 379 VPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 424
Db 301 VPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 346

RESULT 10
US-10-288-798-16

; Sequence 16, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannie B;
; APPLICANT: WALIA, Navinder K.; HAPALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.; Y. Tom;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARZANI, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyoung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Javalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 7477062CD1
US-10-288-798-16

Query Match 24.4%; Score 546.5; DB 12; Length 348;
Best Local Similarity 42.9%; Pred. No. 5e-32;
Matches 117; Conservative 42; Mismatches 111; Indels 3; Gaps 3;
QY 34 EDQALTLRLTASDVTKHYVELRELKGYKVDLVVYKGTGTMALKFVNKSKTKLKN 93
Db 43 EDMTLSAQTLVRAEVDELVEYRPLGQGYRVLVTHRQKGTPLAKQLPKPTSLRG 102
QY 94 FLREVSTNSLSPPFIKVFVDFVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKR 153
Db 103 FLYEFVGLSLGSAISLVTAAGTGISAHYSFLTEPVLHGDLMAFQVGLPQPAVHR 162
QY 154 CVQOGLGLDFHGRQLVHRDIKPEVNLVDFRECRVKLADFGMTRVRCRVKRVSGTIP 213
Db 163 CAAQALASALEYIHARGVLVDRDLKPEVNLVCDPACRRFKLTDFGHTPRGTLRLAGPPI 222
QY 214 YTAPEVC-QAGRADGLAVDTGVVDVWAFGLVLCVLTGNFPWE-AASGADAFFEFVWQ 271
Db 223 YTAPELCAPPLPEGLPIQPALDANALGVLLFCVLTGYFPWDRPLAEDFPYDFDLTWO 282
QY 272 GRPL-GLPSQWRRFTEPALRMFORLLALEPERR 303
Db 283 SGQPRDRPQFWFGLAAAAADALLRGLLDHPHRRR 315

	;	TITLE OF INVENTION:	THEOREF	
	;	FILE REFERENCE:	CL000536-CIP	
	;	CURRENT APPLICATION NUMBER:	US/09/816,094	
	;	CURRENT FILING DATE:	2001-03-26	
	;	NUMBER OF SEQ ID NOS:	4	
	;	SOFTWARE:	FastSEQ for Windows Version 4.0	
	;	SEQ ID NO 4		
	;	LENGTH:	417	
	;	TYPE:	PRT	
	;	ORGANISM:	Rattus norvegicus	
	;		US-09-816-094-4	
		Query Match	93.1%; Score 2084.5; DB 9; Length 417;	
		Best Local Similarity	94.6%; Pred. No. 4.2e-144;	
		Matches	401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;	
QY	1	MSVGCPEPPSRSTCCGPGTAPGGAGVPLLTEDMQALTTLRTLAASDVTKHVELVRELG	60	
DB	1	MSVGCPEPEPLHSLPCCGPGAAPVGAGVPLLTEDMQALTTLRTLAASDVTKHVELVRELG	60	
QY	61	KGTGVKVDLVVKYKTGMTWALKFNYSKTKLKNFLREVSITNSLSSSSPFIKKVDFVPET	120	
DB	61	KGYTGKVDLVVAIKGTGMTWALKFNYSKTKLKNFLREVSITNSLSSSSPFIKKVDFVPET	120	
QY	121	EDCVVFAQEYAPAGDLFDIIPPQVGLPEDTVKRCVCQQQLGLALDFMHGRQLVHRDIKPENV	180	
DB	121	EECVVFAQEYAPAGDLFDIIPPQVGLPEDTVKRCVCQQQLGLALDFMHSRQLVHRDIKPENV	180	
QY	181	LLEDRECRVKLADFGMTTRRVGCRVKRGSIPIPTAPEVCOAGRADGLAVDTGDVWFAG	240	
DB	181	LLEDRECRVKLADFGMTTRRVGCRVKRGSIPIPTAPEVCOAGRADGFANDTGDVWFAG	240	
QY	241	VLIFCVLITGNFPWEAASGADAFFEEFVRWQRGLPLPSQWRRTTEPALRMFORLLALEP	300	
DB	241	VLIFCVLITGNFPWEAASGADAFFEEFVRWQRGLPLPSQWRRTTEPALRMFORLLALEP	300	
QY	301	BRGPAKEVFRFLKHETLSELRPRSHARKPPGDRPPAAPGLRLIEAPGLKRTLVTESG	360	
DB	301	ERRGPAKEVFRFLKHETLSELRPRSHARKPPGDRLP--GPLLEAPGPKRTLVTESG	358	
QY	361	SGRPAPPAPGVSVPLPVPPVPPVPPVPPBGLAPOCGRGRTDGRADKSGQVVLATAI	420	
DB	359	SGKRSPSPSGVPV-----VPVPPVPPVPPBGLAPPAPPGRTDGRADKSGQVVLATAI	413	
QY	421	EICV 424		
DB	414	EICV 417		
		RESULT 7		
		US-09-734-032-4		
		; Sequence 4, Application US/09734032		
		; Patent No. US20020103116A1		
		; GENERAL INFORMATION:		
		; APPLICANT: WEI et al		
		; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC		
		; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES		
		; TITLE OF INVENTION: THEREOF		
		; FILE REFERENCE: CL00536		
		; CURRENT APPLICATION NUMBER: US/09/734,032		
		; CURRENT FILING DATE: 2001-08-16		
		; PRIOR APPLICATION NUMBER: 60205228		
		; PRIOR FILING DATE: 2000-05-17		
		; NUMBER OF SEQ ID NOS: 4		
		; SOFTWARE: FastSEQ for Windows Version 4.0		
		; SEQ ID NO 4		
		LENGTH: 417		
		TYPE: PRT		
		ORGANISM: rat		
		US-09-734-032-4		
		Query Match	93.1%; Score 2084.5; DB 10; Length 417;	
		Best Local Similarity	94.6%; Pred. No. 4.2e-144;	

```
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536DIV
; CURRENT APPLICATION NUMBER: US/10/233,613
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-613-2

Query Match
Best Local Similarity 100.0%; Score 2239; DB 15; Length 424;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
DB 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTQWALKFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120
DB 61 KGTGKVDLVVYKGTGTQWALKFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120
QY 121 EDCYVFAOEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
DB 121 EDCYVFAOEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
QY 181 LFDRECRRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
DB 181 LFDRECRRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVPEVPLPAGPLRLEAPGLKRTVLTESG 300
DB 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVPEVPLPAGPLRLEAPGLKRTVLTESG 300
QY 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
DB 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
QY 361 SGRPAPAVGSVPLPVVPVVPVPEVPLPAGPLRLEAPGLKRTVLTESG 420
DB 361 SGRPAPAVGSVPLPVVPVVPVPEVPLPAGPLRLEAPGLKRTVLTESG 420
QY 421 EICV 424
DB 421 EICV 424

RESULT 5
US-10-288-798-1
; Sequence 1, Application US/10288798
; Publication No: US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B.;
; APPLICANT: WALIA, Narinder K.; HAPALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Javalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BUREFORD, Neil;
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
```

```
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 7312543CD1
US-10-288-798-1
```

```
Query Match
Best Local Similarity 99.8%; Score 2235; DB 12; Length 424;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
DB 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTQWALKFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120
DB 61 KGTGKVDLVVYKGTGTQWALKFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120
QY 121 EDCYVFAOEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
DB 121 EDCYVFAOEYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
QY 181 LFDRECRRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
DB 181 LFDRECRRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVPEVPLPAGPLRLEAPGLKRTVLTESG 300
DB 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVPEVPLPAGPLRLEAPGLKRTVLTESG 300
QY 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
DB 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
QY 361 SGRPAPAVGSVPLPVVPVVPVPEVPLPAGPLRLEAPGLKRTVLTESG 420
DB 361 SGRPAPAVGSVPLPVVPVVPVPEVPLPAGPLRLEAPGLKRTVLTESG 420
QY 421 EICV 424
DB 421 EICV 424
```

```
RESULT 6
US-09-816-094-4
; Sequence 4, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:57:52 ; Search time 38 seconds
(without alignments)
2075.186 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPPRSITCCPGP.....GRADSKGVVLAIEICV 424

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US02_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2239	100.0	424	9	US-09-816-094-2	Sequence 2, Appli
2	2239	100.0	424	10	US-09-734-032-2	Sequence 2, Appli
3	2239	100.0	424	14	US-10-016-985-2	Sequence 2, Appli
4	2239	100.0	424	15	US-10-233-613-2	Sequence 2, Appli
5	2235	99.8	424	12	US-10-288-798-1	Sequence 1, Appli
6	2084.5	93.1	417	9	US-09-816-094-4	Sequence 4, Appli
7	2084.5	93.1	417	10	US-09-734-032-4	Sequence 4, Appli
8	2084.5	93.1	417	15	US-10-233-613-4	Sequence 4, Appli
9	1822	81.4	346	9	US-09-916-790-2	Sequence 2, Appli
10	546.5	24.4	348	12	US-10-288-798-16	Sequence 16, Appli
11	546.5	24.4	348	12	US-10-353-690-50	Sequence 50, Appli
12	394	17.6	256	12	US-09-863-776-40	Sequence 40, Appli
13	390	17.4	256	11	US-09-823-187-96	Sequence 96, Appli
14	383.5	17.1	250	12	US-09-976-782-31	Sequence 31, Appli
15	350	15.6	254	12	US-09-976-782-30	Sequence 30, Appli

16	350	15.6	256	12	US-09-863-776-41	Sequence 41, Appli
17	349.5	15.6	252	12	US-09-976-782-41	Sequence 41, Appli
18	346	15.5	648	15	US-10-024-036B-5	Sequence 5, Appli
19	346	15.5	817	10	US-09-992-481-4	Sequence 4, Appli
20	346	15.5	817	12	US-10-434-034-4	Sequence 4, Appli
21	345.5	15.4	514	12	US-09-949-029-115	Sequence 115, App
22	345.5	15.4	514	12	US-09-949-029-116	Sequence 116, App
23	341.5	15.3	729	12	US-10-161-565-26	Sequence 26, Appli
24	341.5	15.3	729	15	US-10-142-356-11	Sequence 11, Appli
25	341.5	15.3	729	15	US-10-195-101-33	Sequence 33, Appli
26	341.5	15.3	752	10	US-09-835-081-2	Sequence 2, Appli
27	340.5	15.2	713	12	US-10-161-565-27	Sequence 27, Appli
28	340.5	15.2	744	10	US-09-835-081-4	Sequence 4, Appli
29	340	15.2	744	14	US-10-020-038-8	Sequence 8, Appli
30	337.5	15.1	744	10	US-09-919-585-3	Sequence 3, Appli
31	336.5	15.0	688	12	US-10-161-565-28	Sequence 28, Appli
32	336.5	15.0	688	12	US-10-161-565-29	Sequence 29, Appli
33	336	15.0	765	10	US-09-975-326-4	Sequence 4, Appli
34	336	15.0	765	12	US-10-462-887-4	Sequence 4, Appli
35	336	15.0	765	15	US-10-217-357-4	Sequence 4, Appli
36	336	15.0	795	10	US-09-919-585-12	Sequence 12, Appli
37	336	15.0	795	12	US-10-161-565-25	Sequence 25, Appli
38	336	15.0	795	15	US-10-142-356-9	Sequence 9, Appli
39	335.5	15.0	766	10	US-09-975-326-2	Sequence 2, Appli
40	335.5	15.0	766	10	US-09-934-406-2	Sequence 2, Appli
41	335.5	15.0	766	12	US-10-170-789-21	Sequence 21, Appli
42	335.5	15.0	766	12	US-10-368-798-14	Sequence 14, Appli
43	335.5	15.0	766	12	US-10-462-887-2	Sequence 2, Appli
44	335.5	15.0	766	15	US-10-217-357-2	Sequence 2, Appli
45	335	15.0	460	9	US-09-740-627-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-09-816-094-2
; Sequence 2, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Human
; US-09-816-094-2

Query Match	100.0%	Score	2239	DB	9	Length	424
Best Local Similarity	100.0%	Pred. No.	2.3e+155				
Matches	424	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MSVGCPEPPRSITCCPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVRELG	60				
Db	1	MSVGCPEPPRSITCCPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVRELG	60				
Qy	61	KGTGKVDLVVYKGTGTHWALKFNKSKTKLKNFLREVSITNSLSSSPFIKVDVWFET	120				
Db	61	KGTGKVDLVVYKGTGTHWALKFNKSKTKLKNFLREVSITNSLSSSPFIKVDVWFET	120				
Qy	121	EDCVFAOEYAPAGDLFDIIPQVGLPDDTVKRCVQQLGLALDFMHGRLVHRIDKPNV	180				
Db	121	EDCVFAOEYAPAGDLFDIIPQVGLPDDTVKRCVQQLGLALDFMHGRLVHRIDKPNV	180				
Qy	181	LLFDRECRVRVLADFGMTFRYGVCRVKRYSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG	240				
Db	181	LLFDRECRVRVLADFGMTFRYGVCRVKRYSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG	240				

Search completed: December 5, 2003, 09:58:52
Job time : 23 secs

QY	41	LRTLAASDVTXHYELVRELKGNFYKVDLVVKGTGTMKALFPVNKSK-----TKLKN	93
Db	1	MSTFRQEDVEDHYEMGEELGSGQFAVRKCOOKGTGMEVAAKFIKRRLLPSSRRRGVSRE	60
QY	94	FUREVSIITNLSLSSPTIIKVFDDVVFTEDCYYFAQYAPAGDLFDIIPQVGLPEPTVKR	153
Db	61	IREFVSLIPEPI-RHPNITIHDFVFNKTDVWILIELVSGGELFDFLAKESLSTEDPATO	118

Db 49 SATDEQPHIGNYRLQKTIKGNFAKVLARHVLTRGVAVKIIDTQINPTSLQKLFREV 108
Qy 99 SITNSLSSPFIHKVDFVETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOOL 158
Db 109 RIMKIL-NHPNIVKLFVI-EETKTLVMEYASGGEVDFYLVAHGRMKKEARAKFQI 166
Qy 159 GLADFMHGRQLVHRDIKPNVLLFDRECRVVKLADFGMTR--VGCVRKVSCTIPIYTA 216
Db 167 VSAVQYCHQKICVHRDLKAENLLD-DAD-MNIIADFGSNEFTVGNKLDTCFSGPPYAA 224
Qy 217 PEVQAGRADGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMQRGLPG 276
Db 225 PELFGKKYDG---PEVDVWSGLVILTVLSGSLPFDGQNLKE--LRERVLRKGYAVPF 278
Qy 277 LPSQWRRTPEPALRMFORLLALEPERRGPAKEVFR 311
Db 279 YMS-----TDCENLLKLLVLPNPKRGSLEQIMK 307

RESULT 10

US-08-817-832B-31
; Sequence 31, Application US/08817832B
; Patent No. 6579691
; GENERAL INFORMATION:
; APPLICANT: MANDELKOM, Eckhard, et al.
; TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: US
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,832B
; FILING DATE: 28-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP95/04258
; FILING DATE: 30-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94 11 7122.5
; FILING DATE: 28-OCT-1994
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-817-832B-31

Query Match 14.9%; Score 333.5; DB 4; Length 779;
Best Local Similarity 26.2%; Pred. No. 2.5e-19;
Matches 107; Conservative 72; Mismatches 138; Indels 91; Gaps 16;

Qy 45 AASDVTKH--YELVRELKGTGKVDLVVYKGTGTMALFKVNSK--TKLKNFLREV 98
Db 35 SATDEQPHIGNYRLQKTIKGNFAKVLARHVLTRGVAVKIIDTQINPTSLQKLFREV 94
Qy 99 SITNSLSSPFIHKVDFVETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOOL 158
Db 95 RIMKIL-NHPNIVKLFVI-EETKTLVMEYASGGEVDFYLVAHGRMKKEARAKFQI 152
Qy 159 GLADFMHGRQLVHRDIKPNVLLFDRECRVVKLADFGMTR--VGCVRKVSCTIPIYTA 216
Db 153 VSAVQYCHQKICVHRDLKAENLLD-DAD-MNIIADFGSNEFTVGNKLDTCFSGPPYAA 224

Qy 217 PEVQAGRADGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMQRGLPG 276
Db 211 PELFGKKYDG---PEVDVWSGLVILTVLSGSLPFDGQNLKELRERSCLR-GKYRVPF 265
Qy 277 LPSQWRRTPEPALRMFORLLALEPERRGPAKEVFR----- 311
Db 266 YMS-----TDCENLLKLLVLPNPKRGSLEQIMKORMMNVGHHEBELKPYSEPELDND 319
Qy 312 -----FLKHELTSELRRRPSHR-----ARKPP---GDRPPAAGPI----- 343
Db 320 AKRIDIMVTMGFARDEINDALVSKYDEVMATVILLGRKPPPEFEGGSLSSGNLCORSRP 379
Qy 344 -----RLKAPGLK--RTVLTE-----SGSGSRPAPPAVGSVLP 376
Db 380 SSDLNNTLSQPAHLKVQRTISANQKORRFSDHAGPSIPPAVSYTKRP 427

RESULT 11

US-08-924-183-2
; Sequence 2, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-924-183-2

Query Match 14.9%; Score 333; DB 3; Length 476;
Best Local Similarity 28.7%; Pred. No. 1.4e-19;
Matches 102; Conservative 62; Mismatches 138; Indels 54; Gaps 13;

Qy 53 YELVRELKGTGKVDLVVYKGTGTMALFKVNSKT--KLKNFLREVSTNSLSSPFI 110
Db 9 WDLVQTLGEGAYGEVQLAVNRITQAVAVKIVDMKRAIDCPQNIKEICI-NKMLSHENV 67
Qy 111 IKVEDVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOGLALDFMGRQL 170
Db 68 VKFYGHRREGHIQYLFL-EYCSGGELEFDRIEPIGMEPEQDAQRFFHQLMAGVYLHGIGI 126
Qy 171 VHRDIKPNVLLFDRECRVVKLADFGMTR--RVGCR--VGRVSGTIPYTAPEVCOAGRA 225
Db 127 THRDIKPNVLLDERRD--NLKISDFGLATVFRHNNRRELLNKKMCGTLPYVAPELLK--RK 182
Qy 226 DGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMQRGLPGLPSQWRRT 285
Db 183 EFHA--EPVDVWSGIVLTAMLAGELFDQPSDS---CQEYSDMKKKTVLNP--WKID 235
Qy 286 EPALRMFORLLALEPERR-----GPAKEVFRFLKHEL 317
Db 236 SAPLALHLKILVETPSARITIPDIKKRWYKNPLNRAKRPATSGMSSSGFSKH-I 294
Qy 318 TSELRRRP----SHRARKPPGDRPPAAGPLRLAPGLKRTVLTESGSGSRPAPP 368
Db 295 HSNLDFSPVNGSSEETVKFSSSQPEPATGLSLWDTGFSNVDKLVQGISFSQPTCP 350

RESULT 12

US-09-488-364-2
; Sequence 2, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-975-326-2

Query Match          15.0%; Score 335.5; DB 4; Length 766;
Best Local Similarity 26.9%; Pred. No. 1.6e-19;
Matches 108; Conservative 60; Mismatches 172; Indels 61; Gaps 13;

QY      42  RTLAASDVTKHYVELRELKGTGYKVVDLVVYKGTTGAKLKFNKSKTKLKNFL--REVS 99
Db      383 RCESTTLLEKYIKIGVIGDGNFAVVKECIDRSTGKEFALKIIDKAKCCGKEHLENEVS 442

QY      100 ITNSLSSSPFIHKVFDPVVFFETEDCYVFAQBYAPAGDLFDIIPPOVGLPEDTVKFCVOQLG 159
Db      443 ILARVKHPNIIMLVEEMETATE--LFVLMLVKGGDLFDAITSTKYTERDGSAMVINLA 500

QY      160 LALDFMHGROLVHRDIKPENULL--PDRECRRVKLADFGMTTRRVGCRVKRVSGTPIPTAP 217
Db      501 NALRYLHLGISIVHRDIKPENLLVCEYPDGTSKLGLDGLATVVGEGPLYTVCGPTTVAP 560

QY      218 EVQCAGRADGLAVDTG---VDVNAFGLIFCVLTGNFPFEAAAGS-ADAFEEFEFVRWQRG 272
Db      561 EI-----IAETGVGLKVDVWAAGVIITYILLCGPPPSRSENNLOEDLFDQIL---AG 608

QY      273 RLPLGLSQMRFTPEPALRMFORLLALEPERRGPAKEVFRF-----LKHELTSE 320
Db      609 KLEFPAPYWNIDTSAKELLISOMLVNVEARCTAGQILSHFWSDDASQENNMQAEVTGK 668

QY      321 LRPRSHRAKKPGDRPPAAGPLRLAEPGLKRIVLTESG-----SGS-----R 364
Db      669 LKQ----HFNNALPKQNSTTTG-----VSVIMNTALDKEQIFCSKHQODSGRPQMEEPIS 719

QY      365 PAPAVGSLVPVFPVFPVVEPVBPBEGLAQCPGPGRKTGD 405
Db      720 PVPSVBEIIPVGEA-VPAITPESPTEPHCFPPAAPGGERAG 759

RESULT 9
US-09-523-849-32
; Sequence 32, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g2052189
US-09-523-849-32

Query Match          14.9%; Score 334; DB 4; Length 793;
Best Local Similarity 31.3%; Pred. No. 2.3e-19;
Matches 86; Conservative 63; Mismatches 102; Indels 24; Gaps 10

QY      45 AASDVTKH---YELVRELKGTGYKVVDLVVYKGTTGAKLKFNKSK---TKLKNFLREV 98

```

```

; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g3089349
US-09-523-849-33

Query Match      15.3%; Score 341.5; DB 4; Length 729;
Best Local Similarity 24.6%; Pred. No. 4.9e-20;
Matches 117; Conservative 75; Mismatches 138; Indels 145; Gaps 18;

QY 37 QALTLRT-----LAASDVTKH---YELVRELKGTGKVDLVVYKGTGTMALK 82
DB 26 QEVTSRTSRGARCENSTASCADEPHIGNVELLKTIGKNFAKVYKLARHILTGREVAIK 85

QY 83 FWNKSK---TKLKNFLREVSTINSLSPPFIKVPDVVFETEDCVFAQEVAPAGDLFDI 139
DB 86 IIDKTQLNPTSLQKLFREVIRMKIL-NHPNIVKLFVI-EKETLYLIMEYASGGEVFDY 143

QY 140 IPPQVGLPDTVKRCVQQLGALDEHWSQLVHRDIEKENVLLFDRECRVKLADFGMT 199
DB 144 LVAHGRMKEKEARSFRQIVSAVQYCHQRIVHRODKAENLL-DAD-MNRIADFGFSN 201

QY 200 R--VCCRKRVSGTTPYTAPEVCQAGRADGLAVDTGVDMVAFGVLIFCVLTGNFPWEAAS 257
DB 202 EFTVGKLDTCGSPYAPAPLFOGKYDG----PEVDVWSLGVILYTLVSGSLPFDGON 257

QY 258 GADAFEFVVRWQRLPGLPQWRFTPEPALRMFORLLALEPERRGPAKEVFR----- 311
DB 258 LKE--LRERVLRGKYRIPFYMS-----TDCENLLKRFVLNPIKRGTLQEIIMKORWINA 309

QY 312 -----FLXHEL-----TS 319
DB 310 GHEDELKPFVPELDDISDQKRIDIMVGMGYSGEIEQESLSKMKYDEITATYLLGRKSS 369

QY 320 EL-----RRRPSHRARKPPGDRP-----PA----- 339
DB 370 ELDASSSSSSNLAKVRPSSDLNNTSQSPHHKVQSVSSSQKORRYSDHAGPAIPGV 429

QY 340 -AGPIRLR---APGPIKRTVLTE-----SGSGSRPAPPVGVPLVPVPVP 382
DB 430 VAYPKRSQTSADGLKEDGISRSKSSGSAVGKGIGIAPASPMIGNASNPNKADIP 484

RESULT 5
US-08-924-183-8
; Sequence 8, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-488-364-8

Query Match      15.2%; Score 340; DB 4; Length 513;
Best Local Similarity 32.0%; Pred. No. 4.2e-20;
Matches 85; Conservative 48; Mismatches 115; Indels 18; Gaps 7;

QY 45 AASDVTKHYELVRELKGTGKVDLVVYKGTGTMALKFVN-KSKTKLKNFLREVSTINS 103
DB 14 ATREVEGWTLAQTLGEGAYGEVKLLINRQTGGCGMKMVDLKKHPDAANSVRKEVCIQK 73

QY 104 LSSSPFIKVPDVVFETEDCVFAQEVAPAGDLFDIIPQVGLPDTVKRCVQQLGLALD 163
DB 74 MLQDKHILRFFGKRSQGSVEYIFL-EYAGGELFDRIEFDVGMPPQHEAQRVFTQLLSGLN 132

QY 164 FMHGRQLVHRDIEKENVLLFDRECRVKLADFGMTTRRVGCRVK-----RVSGTIPYTAPE 218
DB 133 YLHQRGIAHRLDKPENLLDDEHD--NVKISDFGMATMFRCKGKERLLDKRCGTLIPYVAPE 190

QY 219 VCQAGRADGLAVDTGVDMVAFGVLIFCVLTGNFPWEAASGADAFEFVVRWQRLPGLP 278
DB 191 VLQKAYQPPA-----DLWSCGVILVTMLAGELPNDQPS---TNCTEFTNWRDNDHWQLQ 242

QY 279 SQWRRFTEPALRMFOR-LLALEPERR 303
DB 243 TPWSKLDTLAISLLRKLKLLATSPGTR 268

RESULT 6
US-09-488-364-8
; Sequence 8, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1013
; CURRENT APPLICATION NUMBER: US/09/488,364
; CURRENT FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-488-364-8

Query Match      15.2%; Score 340; DB 4; Length 513;
Best Local Similarity 32.0%; Pred. No. 4.2e-20;
Matches 85; Conservative 48; Mismatches 115; Indels 18; Gaps 7;

QY 45 AASDVTKHYELVRELKGTGKVDLVVYKGTGTMALKFVN-KSKTKLKNFLREVSTINS 103
DB 14 ATREVEGWTLAQTLGEGAYGEVKLLINRQTGGCGMKMVDLKKHPDAANSVRKEVCIQK 73

QY 104 LSSSPFIKVPDVVFETEDCVFAQEVAPAGDLFDIIPQVGLPDTVKRCVQQLGLALD 163
DB 74 MLQDKHILRFFGKRSQGSVEYIFL-EYAGGELFDRIEFDVGMPPQHEAQRVFTQLLSGLN 132

QY 164 FMHGRQLVHRDIEKENVLLFDRECRVKLADFGMTTRRVGCRVK-----RVSGTIPYTAPE 218
DB 133 YLHQRGIAHRLDKPENLLDDEHD--NVKISDFGMATMFRCKGKERLLDKRCGTLIPYVAPE 190

QY 219 VCQAGRADGLAVDTGVDMVAFGVLIFCVLTGNFPWEAASGADAFEFVVRWQRLPGLP 278
DB 191 VLQKAYQPPA-----DLWSCGVILVTMLAGELPNDQPS---TNCTEFTNWRDNDHWQLQ 242

QY 279 SQWRRFTEPALRMFOR-LLALEPERR 303
DB 243 TPWSKLDTLAISLLRKLKLLATSPGTR 268

```

Db 361 SGRPAPPAGVPLPVFPVFPVFPVPEPGLAPOGPRTDGRADKSGQWLATAI 420
 QY 421 EICV 424
 ||||
 Db 421 EICV 424

RESULT 2

US-09-816-094-4
 ; Sequence 4, Application US/09816094
 ; Patent No. 6534299
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL000536-CIP
 ; CURRENT APPLICATION NUMBER: US/09/816,094
 ; CURRENT FILING DATE: 2001-03-26
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 417
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-816-094-4

Query Match 93.1%; Score 2084.5; DB 4; Length 417;
 Best Local Similarity 94.6%; Pred. No. 3e-164;
 Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

QY 1 MSVGPEPEPPESLCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
 Db 1 MSVGPEPEPLSLCCGPGAPVPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
 QY 61 KGYGKVDLVVYKGTGTAKLXFNKSKTKLKNFLREVSITNSLSSPPPIIKVFDVFEET 120
 Db 61 KGYGKVDLVVYKGTGTAKLXFNKSKTKLKNFLREVSITNSLSSPPPIIKVFDVFEET 120
 QY 121 EDCYFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLADFMHGRQLVHRDIKPENV 180
 Db 121 EDCYFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLADFMHGRQLVHRDIKPENV 180
 QY 181 LFDRECRVKLADFGMTREVRGCRVKRVSSTIPTYAPEVCQAGRADGLAVDTGVDVWAFG 240
 Db 181 LFDRECRVKLADFGMTREVRGCRVKRVSSTIPTYAPEVCQAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTPEPALRMFORLLALEP 300
 Db 241 VLIFCVLTGNFPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTPEPALRMFORLLALEP 300
 QY 301 ERGPAKEVERFLKHELTSELRRPDSHRARKPPGDRPAAAGPLRLAEPGLKRTVLTESG 360
 Db 301 ERGPAKEVERFLKHELTSELRRPDSHRARKPPGDRPAAAGPLRLAEPGLKRTVLTESG 360
 QY 361 SGRPAPPAGVPLPVFPVFPVFPVPEPGLAQGPGRDGRADKSGQWLATAI 420
 Db 359 SGRPAPPAGVPLPVFPVFPVFPVPEPGLAQGPGRDGRADKSGQWLATAI 413
 QY 421 EICV 424
 ||||
 Db 414 EICV 417

RESULT 3

US-08-677-298-2
 ; Sequence 2, Application US/08677298
 ; Patent No. 5863729
 ; GENERAL INFORMATION:
 ; APPLICANT: Pivnicka-Worms, Helen
 ; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAK-1
 ; TITLE OF INVENTION: KINASE
 ; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: CO
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/677,298
 ; FILING DATE: 09-JUL-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Caruthers, Jennie M.
 ; REGISTRATION NUMBER: 34,464
 ; REFERENCE/DOCKET NUMBER: 9-96
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303) 499-8080
 ; TELEFAX: (303) 499-8089
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 729 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-677-298-2

Query Match 15.3%; Score 341.5; DB 2; Length 729;
 Best Local Similarity 24.6%; Pred. No. 4.9e-20;
 Matches 117; Conservative 75; Mismatches 138; Indels 145; Gaps 18;

QY 37 QALTLRT-----LAASDVTKH---YELVRELKGYTKGVLDLVVYKGTGTAKMALK 82
 Db 26 QEVTSRTSRSGARCNSTASCADQPHGNVRLTKTGKGNFAKVLARHLITGREVAIK 85
 QY 83 FVNSK---TKLKNFLREVSTNSLSSPFIKVPDVFVETEDCVFAQEVAPAGDLFDI 139
 Db 86 IIDQTQNPSTLSLOKLFREVRIMKIL-NHPNIVKLFVI-EETKTLYLIMEYASGGEVFDY 143
 QY 140 IPPQVGLPDTVKRCVQQLGLADFMHGRQLVHRDIKPENVLLFPRECRVKLADFGMTR 199
 Db 144 LVAHGRMEKEARSKFRQIVSAVOYCHOKRIVHRDLKAENLL-DAD-WNTIKIADFGFSN 201
 QY 200 R--VGCRCVKRVSSTIPTYAPEVCQAGRADGLAVDTGVDVWAFGVLIFCVLTGNFPWEAAS 257
 Db 202 EFTVGKLDFTFCGSPPYAAPLFOGKKYDG----PEVDVMSLGVILYTLVSGSLPFDGQN 257
 QY 258 GADAFEEFVRWQGRGLPGLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFR----- 311
 Db 258 LKE--LRERVLRGKVRIPFYS-----TDCENLLKRFVLNPIKRGTLQEQIMKORWINA 309
 QY 312 -----FLXHEL-----TS 319
 Db 310 GHEEDELKPFVEPELIDISQKIDIMVGMYSQEBIQESLSKMKYDEITATVLLGRKSS 369
 QY 320 EL-----RRRPSHRARKEPGDRP-----PA----- 339
 Db 370 ELDASSSSSSNLAKVTPSSDLNNS--QSPHHKVRQSVSSSQKORRYSDHAGPAIPSV 429
 QY 340 -AGPLRLE---APGPLKRTVLTE-----SGSGSRPAPPAGVPLPVFPVFPV 382
 Db 430 VAYPKESQTSADGDLKEDGISRRKSSGSAGVGKGIAPASPMIGNASNPNKADIP 484

RESULT 4

US-09-523-849-33
 ; Sequence 33, Application US/09523849
 ; Patent No. 6458561
 ; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 Seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPEPRLTCCGPG.....GRADKSGQVVLATAFICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents:AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*

5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	4	US-09-816-094-2
2	2084.5	93.1	417	4	US-09-816-094-4
3	341.5	15.3	729	2	US-08-677-298-2
4	341.5	15.3	729	4	US-09-523-849-33
5	340	15.2	513	3	US-08-924-183-8
6	340	15.2	513	4	US-09-488-364-8
7	336	15.0	765	4	US-09-975-326-4
8	335.5	15.0	766	4	US-09-975-326-2
9	334	14.9	793	4	US-09-523-849-32
10	333.5	14.9	779	3	US-08-817-832B-31
11	333	14.9	476	3	US-08-924-183-2
12	333	14.9	476	4	US-09-488-364-2
13	328.5	14.7	448	2	US-09-159-385-2
14	328.5	14.7	448	3	US-09-186-277-2
15	328	14.6	476	3	US-09-522-800-16
16	328	14.6	476	3	US-08-924-183-1
17	328	14.6	476	4	US-09-488-364-1
18	328	14.6	476	4	US-09-113-785-1
19	323.5	14.4	722	4	US-09-984-890-4
20	323.5	14.4	722	4	US-08-817-832B-32
21	322.5	14.4	483	3	US-08-924-183-9
22	322.5	14.4	483	4	US-09-488-364-9
23	322.5	14.4	745	4	US-09-523-849-36
24	321.5	14.4	724	4	US-09-984-890-2
25	318	14.2	454	2	US-09-159-385-1
26	318	14.2	454	3	US-09-186-277-1
27	315	14.1	552	3	US-08-557-006C-40

28	311.5	13.9	582	4	US-09-428-711A-2	Sequence 2, Appli
29	311.5	13.9	1051	4	US-09-428-711A-14	Sequence 14, Appli
30	310	13.8	668	4	US-09-930-181-2	Sequence 2, Appli
31	309	13.8	633	3	US-08-557-006C-43	Sequence 43, Appli
32	308.5	13.8	257	2	US-07-857-224B-25	Sequence 25, Appli
33	307	13.7	543	4	US-09-529-093A-2	Sequence 2, Appli
34	307	13.7	543	4	US-09-529-154-2	Sequence 7, Appli
35	305.5	13.6	476	3	US-08-924-183-7	Sequence 7, Appli
36	305.5	13.6	476	3	US-09-488-364-7	Sequence 11, Appli
37	302	13.5	631	4	US-09-523-849-11	Sequence 34, Appli
38	302	13.5	776	4	US-09-523-849-34	Sequence 1, Appli
39	298.5	13.3	345	3	US-09-101-146-1	Sequence 35, Appli
40	298	13.3	604	4	US-09-523-849-35	Sequence 6, Appli
41	295.5	13.2	257	3	US-09-101-146-6	Sequence 2, Appli
42	294.5	13.2	396	4	US-09-801-876B-2	Sequence 16, Appli
43	294.5	13.2	1050	4	US-09-428-711A-16	Sequence 2, Appli
44	294	13.1	436	4	US-09-734-673-2	Sequence 2, Appli
45	294	13.1	436	4	US-09-523-849-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-816-094-2
; Sequence 2, Application US/09816094

; Patent No. 6534299

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CLO00536-CIP

; CURRENT APPLICATION NUMBER: US/09/816,094

; CURRENT FILING DATE: 2001-03-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Human

US-09-816-094-2

Query Match 100.0%; Score 2239; DB 4; Length 424;
Best Local Similarity 100.0%; Pred. No. 5.4e-177; Indels 0; Gaps 0;
Matches 424; Conservative 0; Mismatches 0;

QY	1	MSVGCPEPEPRLTCCGPGTAPGPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG	60
Db	1	MSVGCPEPEPRLTCCGPGTAPGPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG	60
QY	61	KGTQKVDLVVYKGTGTMALFKVNSKTKLKNFLREVSITNSLSSSPFIIVKVDVVFET	120
Db	61	KGTQKVDLVVYKGTGTMALFKVNSKTKLKNFLREVSITNSLSSSPFIIVKVDVVFET	120
QY	121	EDCYFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMGRQLVHRDIKPNV	180
Db	121	EDCYFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMGRQLVHRDIKPNV	180
QY	181	LLFDRCRRVKLADFGMTRRVGRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG	240
Db	181	LLFDRCRRVKLADFGMTRRVGRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG	240
QY	241	VLIFCVLTGNFWEAASGADAFEEFVRWQRGLPGLPSQWRFFTEPALRMFORLLALBP	300
Db	241	VLIFCVLTGNFWEAASGADAFEEFVRWQRGLPGLPSQWRFFTEPALRMFORLLALBP	300
QY	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPAAGPLRLLEAPGLKRTVLTESG	360
Db	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPAAGPLRLLEAPGLKRTVLTESG	360
QY	361	SGSRPAPPVAVGSVPLPVFPVFPVPEPGIAPQGPPTDGRADKSGQVVLATAI	420

5/17/00
60205228

Db	1386	GGCAGGACCGACGCGCGTGCAGACAAAGCGCAGGTGGTATTGGCCACAGCCATC	1445
Qy	1261	GAGATCTCGCTCTGA	1275
Db	1446	GAGATCTCGCTCTGA	1460
RESULT 15			
AX589475	AX589475	936 bp	DNA
LOCUS	AX589475	Sequence 4 from Patent WO02083882.	linear PAT 24-JAN-2003
DEFINITION	AX589475	AX589475	
ACCESSION	AX589475.1	GI:27900911	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT	155 a	308 c	308 g 165 t
ORIGIN			
Query Match		73.0%;	Score 931.2; DB 6; Length 936;
Best Local Similarity		99.7%;	Pred. No. 2.7e-109;
Matches	933;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	1	ATGAGCGTGGGTGCGCAGAGCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGGCGCGGG	60
Db	1	ATGAGCGTGGGTGCGCAGAGCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGGCGCGGG	60
Qy	61	ACTGCCCCCTGGCCCTGGTCCCGGTGCCCCCTTCTCACTGAAGACATGCAGGCCCTGACT	120
Db	61	ACTGCCCCCTGGCCCTGGTCCCGGTGCCCCCTTCTCACTGAAGACATGCAGGCCCTGACT	120
Qy	121	CTCCGCACACTGCGCGCCACAGCAGCTCACAAAGCACTAGCAACTATCGGGAGCTGGGC	180
Db	121	CTCCGCACACTGCGCGCCACAGCAGCTCACAAAGCACTAGCAACTATCGGGAGCTGGGC	180
Qy	181	AAAGGACCTATGGAGAGTTGACCTGGTGGTCTACAGGGCACAGGCACAAATGGCA	240
Db	181	AAAGGACCTATGGAGAGTTGACCTGGTGGTCTACAGGGCACAGGCACAAATGGCA	240
Qy	241	CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC	300
Db	241	CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC	300
Qy	301	ACCAACAGCCTCTCCTCCAGCCCTTTCATCAAGGTCTTTGACGTCTTTTGAGACA	360
Db	301	ACCAACAGCCTCTCCTCCAGCCCTTTCATCAAGGTCTTTGACGTCTTTTGAGACA	360
Qy	361	GAGGACTGTACGTCTTTTCCAGGAGTAGCGACCTGCTGGGACCTGTTGACATCATC	420
Db	361	GAGGACTGTACGTCTTTTCCAGGAGTAGCGACCTGCTGGGACCTGTTGACATCATC	420
Qy	421	CTTCCCCAGGTGGGCTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCTG	480
Db	421	CTTCCCCAGGTGGGCTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCTG	480
Qy	481	GCCTGGACTTTCATGACGGCGGCGAGCTGGTGCACCGGACATCAAGCCCGAGAACGTG	540
Db	481	GCCTGGACTTTCATGACGGCGGCGAGCTGGTGCACCGGACATCAAGCCCGAGAACGTG	540
Qy	541	CTGCTGTTGACCGGAGGTGCCCGCGGTAAAGCTGGCGACTTTCGGCATGACCGCGCG	600

Search completed: December 7, 2003, 16:12:00
Job time : 3236 secs

REMARK	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: gcaps-r@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 37 Row: f Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. Location/Qualifiers 1. 4123 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clones="MGC:28537 IMAGE:4194152" /tissue_type="Liver, normal. 5 month old male mouse." /clone_lib="NCI CGAP_Li9" /lab_host="DH10B" /note="vector: PCMV-SPORT6" 1. 4123 /gene="Sbk-pending" /db_xref="LocusID:104175" /db_xref="MGI:2135937" 489..1742 /codon_start=1 /product="SH3-binding kinase" /protein_id="AAH31759.1" /db_xref="GI:21594544" /db_xref="LocusID:104175" /translation="MSVCSPEPELHSLPCCGPGAAVPVAGVPLLTEDMOALTILTL AASVTKYELVRELKGTGYKVDLVAYKGTGKMAKFNKSKTKLANEIVSITN SLSSPFLIKVDFVFEIECTVFAQETAPAGDLFIIPQVGLPDTVKSCVQLGL ALDFMHSQLVHDKIPENLVLLFDECRRLADFGMTTRVRCVRVSGRIIPIAPE VQAGRADGFVDTGDVMAFGVLIFCVLTGNFWEAAGADAFEEFVRWQRGLPG LPSQWRRTPEPALRFORLLALEPERRGPAPKEVFLKHELTSELRSPSHRARKPG DRLPSLRLEAPGLKRVLTNLSGSGSRSPSPVGVVPVVPVPEAGLAPPAP PGRDGRDTSKSGOVLTALTEICV"
gene	BASE COUNT 838 a 1263 c 1205 g 817 t
CDS	ORIGIN Query Match 75.8%; Score 966.6; DB 10; Length 4123; Best Local Similarity 86.3%; Pred. No. 6e-114; Matches 1100; Conservative 0; Mismatches 154; Indels 21; Gaps 2; QY 1 ATGAGCGTGGGTGCCAGAGCTAGCGCGCCCGCTCCCTACCTGCTGTGGCGCGGG 60 Db 489 ATGAGCGTGGGTGCCAGAGCTAGCGCGCCCGCTCCCTACCTGCTGTGGCGCGGG 548 QY 61 ACTGCCCTGGGCTGGTCCGCTGGTGGCCCTTCTCACTGAGACATGACGCGCTGACT 120 Db 549 GCGGCCCTGTACAGGTGAGGTGTGGCCCTCTCTCACAGACATGACAGCGCTGACC 608 QY 121 CTCGCGACACTGCGCGCGACGACGTCACCAAGCACTACCACTAGTCGCGGAGCTGGGC 180 Db 609 CTACGGACACTGCTGCCAGCGATGTTACCAAGCACTATGAGCTTGTGGGAGCTGGGC 668 QY 181 AAAGGCACTATGGGAAGTTGACCTGGTGGTCTAAGGGGCAAGGCAAAAAATGGCA 240

/db xref="GI:19388002"		810 a 1193 c 1145 g 805 t	
/db xref="LocusID:104175"			
/translation="MSVGPPEPEPLSLPCCGPGKALPVPAGVPLLTEDMQLTERTL			
AASDVKHVELRELKGTGKVDLYVAGTGTGKMAKFKVNSKTLKXNPLREVSITN			
SLSSPFIKVPDVPETECYVFAQEPAGLFDIIPVOGLPEDTVKRCVQOGL			
ALDFMSRQIUVHRIKDPENVILLFDRECRVKLADFGMTSRVVGCRVKSVTIPTYAPE			
VCQAGRADGPAVDITGVDFVAFGLIIFCVLTGNFWEASADAFEFVFNQRGRIFG			
LPSQWRFTFPAIRLQRLALPEPRRGPAPKVFRLKHELTSELRRRPSHRARKPFG			
DRLPGLRLAPGLPKETVLTESGSRSPSPVGPVVPVFPVPEPAGLAPP			
PGRTDGETDKSGQVLAIAIECV"			
BASE COUNT		810 a 1193 c 1145 g 805 t	
ORIGIN			
Query Match		75.8%; Score 966.6; DB 10; Length 3953;	
Best Local Similarity		86.3%; Pred. No. 6.1e-114;	
Matches 1100; Conservative		0; Mismatches 154; Indels 21; Gaps 2;	
Qy	1	ATGAGCGTGGCTGCCAGACCTGAGCCGCGCCCTCCCTGACCTGCTGTGGGCGGGG	60
Db	326	ATGAGCGTGGCTGCCAGACCTGAAACCGCTCCACTCCCTGCTGTGGGCGGGG	385
Qy	61	ACTGCCCTGGCTGGTGGCGGTGTGCCCTTCTCACTGAAGACATGACGCGCTGACT	120
Db	386	GCCGCCCTGTACAGGTGACGTGTGCCCTTCTCACTGAAGACATGACGCGCTGACC	445
Qy	121	CTCCGCACTGGCGCCGACGACGTACCAAGCACTACGAACCTAGTCCGGGAGCTGGC	180
Db	446	CTACGCACTGGCTGGCGAGCATGTTACCAAGCACTATGAGCTGTGCGGGAGCTGGC	505
Qy	181	AAAGGCACTATGGAAGTTGACTGTGTGTCTACAGGSCACAGGCACAAATGCA	240
Db	506	AAAGGCACTACGGAAGCTGCACTGTGTGTCTACAGGSCACAGGCACAAATGCCC	565
Qy	241	CTGAAGTTGTGAACAAGCAAAACCAAGCTGAAAGAACTTCTTCAAGGAGTGAGCATC	300
Db	566	CTGAATTTGTGAATAAGATGAGCAAGCTGAAAGAACTTCTTCAAGGAGTGAGCATC	625
Qy	301	ACCAACAGCTTCTCTCCAGCCCTTCAATCAATCAAGGCTTTGACGTGTCTTTGAGACA	360
Db	626	ACCAACAGCTTCTCTTAGCCCTTCAATCAATCAAGGCTTTGACGTGTCTTTGAGACC	685
Qy	361	GAGGCTGCTAGCTTTTCCAGAGTACGACCTGCTGGGACCTTTTACATCATC	420
Db	686	GAGGAGTGCTAGCTTTTCCAGAGTATGACCTGCTGGGACCTTTTACATCATC	745
Qy	421	CTTCCCAAGTGGCTCTCTGAGCACACGCTGAAAGCGCTGTGTGACGAGCTGGGCTG	480
Db	746	CTTCTCTAGTGGGCTCTCCAGAGCACACGCTGAAAGCGCTGTGTGACGAGCTGGGCTG	805
Qy	481	GCGCTGGACTTCATGACGGGCGGAGCTGGTGTGACCGGACATCAAGCCGAGAACGTG	540
Db	806	GCGCTGGACTTCATGACAGCAGGAGCTGGTGTGACCGGACATCAAGCCGAGAACGTG	865
Qy	541	CTGCTGTTGACCGGAGTGGCGCGCTTAAAGCTGGCGACTTGGGATGACGCGCGC	600
Db	866	CTGCTGTTGACCGGAGTGGCGCGCTTAAAGCTGGCGACTTGGGATGACGCGCGC	925
Qy	601	GTGGGCTGGCGCTCAAGCGCGTGAAGCGCACCATCCCTTACACGCGCGCTGAGGTGTG	660
Db	926	GTGGGCTGGCGCTGAGCGCGTGAAGCGCACCATTAACCTACACAGCACCCGAGGTGTG	985
Qy	661	CAGGCGGCGCGCGCGAGCGGTGGCGGTGGACACGCGGCTGACCTGTGGCGCTTCGCG	720
Db	986	CAGGCTGGCGCGCGAGCGGTGGCGGTGGACACGCGGCTGACCTGTGGCGCTTCGCG	1045
Qy	721	GTGCTCATCTTCTGCTGCTACCGCAACTTCCCTGGAGGCGGCTCGGCGCGCGAC	780
Db	1046	GTGCTCATCTTCTGCTGCTACCTGCAACTTCCCTGGAGGCGGCTGCATCGGCTGGGAT	1105
Qy	781	GCCTTCTTGAGAGTGTGTGCTGGCAGCGGGCGCGCTCGCGGGGCTTCCTTCGACG	840
Db	1106	GCCTTCTTGAGAGTGTGTGCTGGCAGCGGGGCTCGCGGGGCTGCATCGCAG	1165

Qy	841	TGGCGCGGCTTACCGAGCGCGCGCTGCGCATTTCCAGCGCTTACTGGCCCTGGAGCC	900
Db	1166	TGGCGCGGCTTACCGAGCGCTGCGCTGCGTATTTCCAGCGCTTCTGGCGCTAGAGCCT	1225
Qy	901	GAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTTCAAGCACGAGCTCACTCGGAG	960
Db	1226	GAGCGCGGTTGGCGCGCGCGCGCGCGCGCTTCTTCAAGCACGAGCTCACTCGGAG	1285
Qy	961	CTGCGCGCGCGCGCGCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1020
Db	1286	CTGCGCGCGCGCGCTTCCACCGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCG	1339
Qy	1021	GGCGCATGCGCTCGAGGCGCGCTGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGC	1080
Db	1340	GGGTCCCTGCGCTTGGGCTTCCGCGCGCTCAAGCGGACGCTGCTGCTACCGAGAGTGC	1399
Qy	1081	AGCGGCTTCCCGCGCGCGCGCGCGCGCGCGCTGCGGTGCGTTCGCGCGCGCGCG	1140
Db	1400	AGCGGCTTCCCGCGCGCTTCCCGCGCGCTAGGCGCGCT-----GGTACCC	1444
Qy	1141	GTGCGAGTGGCG	1200
Db	1445	GTGCGAGTGGCGCGCTTCCAGTACCGCGCGCTGAGGCTGTGCGCTGCGCGCGCG	1504
Qy	1201	GGCGGACCGCATC	1260
Db	1505	GGCGGACCGCATC	1564
Qy	1261	GAGATCTGCGCTCTGA	1275
Db	1565	GAGATCTGCGCTCTGA	1579
RESULT 12			
BC031759			
LOCUS			
DEFINITION			
IMAGE:4194152), complete cds.			
ACCESSION			
BC031759.1 GI:21594543			
KEYWORDS			
MGC.			
SOURCE			
Mus musculus (house mouse)			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 4123)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Schaeetz, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S.,			
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
Abramsen, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,			
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Fahey, J., Hulton, E., Kettman, M., Madan, A., Rodriguez, S.,			
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzyzanski, M.I., Skalek, U., Smalls, D.E.,			
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
JOURNAL			
MEDLINE			
PUBMED			
22388257			
12477932			
2 (bases 1 to 4123)			
Strausberg, R.			
Direct Submission			
TITLE			
Submitted (06-JUN-2002) National Institutes of Health, Mammalian			
JOURNAL			

```

650 CCTCTCAGTGGGCTCCAGAGGACACGCTGAAGCGCTGTGTGCAGCAGCTGGGCGTG 709
481 GCGCTGACATTCATGACGCGGCGCAGCTGTGTGCACCGACATCAAGCCGAGACGCTG 540
710 GCGCTGACATTCATGACGCGGCGCAGCTGTGTGCACCGACATCAAGCCGAGACGCTG 769
541 CTGCTGTTCACGCGGAGTGGCGCGGTAAAGCTGGCCGACATTCGGCATGACGCGCGC 600
770 CTGCTGTTCACGCGGAGTGGCGCGGTAAAGCTGGCCGACATTCGGCATGACGCGCGC 829
601 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACACATTCCTTACAGCGCGCTGAGTGTGC 660
830 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACACATTCCTTACAGCGCGCTGAGTGTGC 889
661 CAGCGGGCGCGCGCGCTGCGGTGACGCGGACAGCGGCTGTGTGCGCTTTCGGCCTTCGG 720
890 CAGCGTGGCGCGCGCGCTTTCGGGTGACAGCGGCTGTGTGCGCTTTCGGCCTTTCGGC 949
721 GTGCTCATCTTCTCGTGTCTCACCGGCAACTTCCCGTGGGAGGCGGCTCGGGCGCCGAC 780
950 GTGCTCATCTTCTCGTGTCTCACCGGCAACTTCCCGTGGGAGGCTGCATCGGTGGGAT 1009
781 GCCTTCTTCGAGGATTCGTGCGTGTGACGCGGCGCGCTTTCGGGTGCGCTTTCGGC 840
1010 GCCTTCTTCGAGGATTCGTGCGTGTGACGCGGCGCGCTTTCGGGTGCGCTTTCGGC 1069
841 TGGGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAGCG 900
1070 TGGGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAGCG 1129
901 GAGCGCGCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAG 960
1130 GAGCGCGCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAG 1189
961 TGGCGCGCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAG 1020
1190 TGGCGCGCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAG 1243
1021 GGGCCACTGCGCTCGAGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAGCGCGCG 1080
1244 GGGTCTCTGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCTTTCAG 1303
1081 AGCGGCTCGCGCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCG 1140
1304 AGCGGCTCGCGCGCTTTCAGCGCGCGCGCTGCGGTGACGCGCTTTCAGCGCGCGCG 1348
1141 GTGCCAGTCCCGTGGCGGCTGCGGTGACGCGCGCGCTGCGGTGACGCGCGCGCGCGCG 1200
1349 GTGCCAGTCCCGTGGCGGCTGCGGTGACGCGCGCGCTGCGGTGACGCGCGCGCGCGCG 1408
1201 GGCAGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
1409 GGCAGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1468
1261 GAGATCTGCGCTGGA 1275
1469 GAGATCTGCGCTGGA 1483

RESULT 11
BC025837
LOCUS BC025837
DEFINITION Mus musculus SH3-binding kinase, mRNA (cdna clone MGC:37964
IMAGE:5135723), complete cds.
ACCESSION BC025837
VERSION BC025837.1
KEYWORDS GI:19388001
SOURCE MGC.
ORGANISM Mus musculus (house mouse)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3953)

```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udín, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Werkley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Soergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Buffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dietlen, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, V.I., Skalska, U., Smalish, D.E.,
Schnier, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

TITLE

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3953)

Strausberg, R.

Direct Submission

Submitted (11-MAR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline

Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 57 Row: n Column: 20

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

Location/Qualifiers

1. 3953

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:37964 IMAGE:5135723"

/tissue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Li9"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 3953

/gene="Shk-pending"

/db_xref="LocusID:104175"

/db_xref="MGI:2135937"

326. 1579

/codon_start=1

/product="SH3-binding kinase"

/protein_id="AAH25837.1"

FEATURES

source

gene

CDS

QY	425	CCAGGTGGGGCTCCCTGTAGGACACGGTGAAGCGCTGTGTGACAGAGCTGGGCTGGCGC	484
Db	464	CCAGGTGGGGCTCCCTGTAGGACACGGTGAAGCGCTGTGTGACAGAGCTGGGCTGGCGC	523
QY	485	TGACTTTATGACAGGGCGGCGAGCTGTGTACCGGACATCAAGCCCGAGAAAGTGTGTC	544
Db	524	TGGACTTATGACAGGGCGGCGAGCTGTGTACCGGACATCAAGCCCGAGAAAGTGTGTC	583
QY	545	TGTTCGACCGGAGTGCAGCGGCTGAAGCTGGCCGACTTTCGGCATGACGCGCGCGTGG	604
Db	584	TGTTTCGACCGGAGTGCAGCGGCTGAAGCTGGCCGACTTTCGGCATGACGCGCGCGTGG	643
QY	605	GCTGCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACAGGGCGCTTGAGGTGTGCCAGG	664
Db	644	GCTGCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACAGGGCGCTTGAGGTGTGCCAGG	703
QY	665	CGGCGCGCGGAGCGGCTGGCTGACACGGGCTGGAGCTGTGGGCTTCGGCGTGC	724
Db	704	CGGCGCGCGGAGCGGCTGGCTGACACGGGCTGGAGCTGTGGGCTTCGGCGTGC	763
QY	725	TCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGGAGGGCGCTCGGGCGCCGACGCT	784
Db	764	TCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGGAGGGCGCTCGGGCGCCGACGCT	823
QY	785	TCCTTCGAGGAGTTCGTGGCTGGCAGCGGGCGCCCTGCGGGGCTGCTTTCGAGTGGC	844
Db	824	TCCTTCGAGGAGTTCGTGGCTGGCAGCGGGCGCCCTGCGGGGCTGCTTTCGAGTGGC	883
QY	845	GCCGCTTACAGACCGCGCTGGCATGTTCCAGCGCTTACTGGCCCTGGAGCCCGAGC	904
Db	884	GCCGCTTACAGACCGCGCTGGCATGTTCCAGCGCTTACTGGCCCTGGAGCCCGAGC	943
QY	905	GCCGCGCGGACCAAGAGGTGTCGCTTCCCTCAAGCAGAGCTCACGTCGAGCTGC	964
Db	944	GCCGCGCGGACCAAGAGGTGTCGCTTCCCTCAAGCAGAGCTCACGTCGAGCTGC	1003
QY	965	GCCGCGCGGCTTCGACACCGCGCGCAAGCGCCCGCGGGGACCGCGCGCGCGCGGCG	1024
Db	1004	GCCGCGCGGCTTCGACACCGCGCGCAAGCGCCCGCGGGGACCGCGCGCGCGCGGCG	1063
QY	1025	CACGCGGCTCAGGCGCTGGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGACGC	1084
Db	1064	CACGCGGCTCAGGCGCTGGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGACGC	1123
QY	1085	GCTCCCGCGCGCGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1144
Db	1124	GCTCCCGCGCGCGCGCGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1183
QY	1145	CAGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1204
Db	1184	CAGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1243
QY	1205	GGACCGAGCGCGCGGACAGAGCAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGG	1264
Db	1244	GGACCGAGCGCGCGGACAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1303
QY	1265	TCTGGCTCTGA	1275
Db	1304	TCTGGCTCTGA	1314
RESULT 9	AX455339		
LOCUS	AX455339	1041 bp	DNA linear PAT 06-JUL-2002
DEFINITION	Sequence 3 from Patent WO0210401.		
ACCESSION	AX455339		
VERSION	AX455339.1		GI:21714512
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		

AUTHORS		Meyers, R., Kapeller-Libermann, R. and Silos-Santiago, I.	
TITLE		18431 and 32374, human protein kinase family members and uses therefor	
JOURNAL		Patent: WO 0210401-A 3 07-FEB-2002; Millennium Pharmaceuticals, Inc. (US)	
FEATURES		Location/Qualifiers	
source		1..1041	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
BASE COUNT		147 a 372 c 359 g 163 t	
ORIGIN			
Query Match		81.6%; Score 1041; DB 6; Length 1041;	
Best Local Similarity		100.0%; Pred. No. 3e-123;	
Matches 1041; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	235	ATGGCACTGAAGTTTGTGAACAGAGCAAAACAAAGCTGAAGAACTTCTTACGGGAGGTG	294
Db	1	ATGGCACTGAAGTTTGTGAACAGAGCAAAACAAAGCTGAAGAACTTCTTACGGGAGGTG	60
QY	295	AGCATCACCAACAGCCTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTCTTT	354
Db	61	AGCATCACCAACAGCCTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTCTTT	120
QY	355	GAGACAGAGACTGTCTTGTCTTGTCCAGGAGTACGACCTGCTGGGACCTGTTTGAC	414
Db	121	GAGACAGAGACTGTCTTGTCTTGTCCAGGAGTACGACCTGCTGGGACCTGTTTGAC	180
QY	415	ATCATCCTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGAGCTG	474
Db	181	ATCATCCTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGAGCTG	240
QY	475	GGCTGCGGCTGATTCATGCAAGGCGGAGCTGTGTGCAACGCGGACATCAAGCCCGAG	534
Db	241	GGCTGCGGCTGATTCATGCAAGGCGGAGCTGTGTGCAACGCGGACATCAAGCCCGAG	300
QY	535	AACGTGCTGTGTGACACCGGAGTGCCTGCGCGTAAAGCTGGCCGACTTCGGCATGACG	594
Db	301	AACGTGCTGTGTGACACCGGAGTGCCTGCGCGTAAAGCTGGCCGACTTCGGCATGACG	360
QY	595	CGCGCGTGGGCTGCGCGCTCAAGCGGCTGAGCGGACCATCCCTTACAGCGGCTCTGAG	654
Db	361	CGCGCGTGGGCTGCGCGCTCAAGCGGCTGAGCGGACCATCCCTTACAGCGGCTCTGAG	420
QY	655	GTGTGCGAGCGGCGCGCGCGAGCGGCTGCGCGTGGACACGGGCTGAGAGCTGTGGGC	714
Db	421	GTGTGCGAGCGGCGCGCGCGAGCGGCTGCGCGTGGACACGGGCTGAGAGCTGTGGGC	480
QY	715	TTGCGGCTGTCTCATCTTCTGCGTGTCTCAACGGGAACTTCCCGTGGGAGCGGCTGGGC	774
Db	481	TTGCGGCTGTCTCATCTTCTGCGTGTCTCAACGGGAACTTCCCGTGGGAGCGGCTGGGC	540
QY	775	GCCGACGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGCGCGCTGCGGGGCTGCCT	834
Db	541	GCCGACGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGCGCGCTGCGGGGCTGCCT	600
QY	835	TGCGAGTGGCGCGCTTCAACGAGCGCGCTGCGCATGTTTCCAGCGCTTACTGGCCCTG	894
Db	601	TGCGAGTGGCGCGCTTCAACGAGCGCGCTGCGCATGTTTCCAGCGCTTACTGGCCCTG	660
QY	895	GAGCCGAGCGCGCGGCGCGGAGGAGTGTTCGCTTCCCTCAAGCAGAGCTCAAG	954
Db	661	GAGCCGAGCGCGCGGCGCGGAGGAGTGTTCGCTTCCCTCAAGCAGAGCTCAAG	720
QY	955	TCCGAGCTGGCGCGCGGCGCTTCGACCGCGCGGCAAGCCCCCGGGGAGCCGCGCGCC	1014
Db	721	TCCGAGCTGGCGCGCGGCGCTTCGACCGCGCGGCAAGCCCCCGGGGAGCCGCGCGCC	780
QY	1015	GCGCGCGGCGCTGCGGCTCGAGGCGCTGGGCGCGCTCAAGCGGAGCGGTGCTGACCGAG	1074
Db	781	GCGCGCGGCGCACTGCGCTTCGAGGCGCTGGGCGCGCTCAAGCGGAGCGGTGCTGACCGAG	840


```
QY 181 AAAGGACCTATGGGAAGTTGACCTGGTGTCTACAAGGGCAGGACCAAAAATGGCA 240
Db 181 AAAGGACCTATGGGAAGTTGACCTGGTGTCTACAAGGGCAGGACCAAAAATGGCA 240
QY 241 CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC 300
Db 241 CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC 300
QY 301 ACCAAGCCTCTCTCCAGCCCTTCAATCATCAAGTCTTTGACGTTTGTGAGACA 360
Db 301 ACCAAGCCTCTCTCCAGCCCTTCAATCATCAAGTCTTTGACGTTTGTGAGACA 360
QY 361 GAGGACTGTACGCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
Db 361 GAGGACTGTACGCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
QY 421 CCTCCCAAGTGGGCTCCCTGAGGACACGGTGAAGCTGTGTGACGACCTGGGCTG 480
Db 421 CCTCCCAAGTGGGCTCCCTGAGGACACGGTGAAGCTGTGTGACGACCTGGGCTG 480
QY 481 CGCTGGACTTCATGACCGGCGGAGCTGTGTGACCGGACATCAAGCCCGGAGACGTG 540
Db 481 CGCTGGACTTCATGACCGGCGGAGCTGTGTGACCGGACATCAAGCCCGGAGACGTG 540
QY 541 CTGCTGTTTCGACCGGAGTGGCGGCTTAAAGCTGGCGGACTTCGGCATGACGGCGC 600
Db 541 CTGCTGTTTCGACCGGAGTGGCGGCTTAAAGCTGGCGGACTTCGGCATGACGGCGC 600
QY 601 GTGGGCTGCCGCGTCAAGCGGTGAGCGGACCATCCCTTACAGGCGCCTGAGGTGTC 660
Db 601 GTGGGCTGCCGCGTCAAGCGGTGAGCGGACCATCCCTTACAGGCGCCTGAGGTGTC 660
QY 661 CAGCGGGCGCGCGGAGTGGCGGTGAGACACGGCGGTGAGCGTGTGGGCTTCGGC 720
Db 661 CAGCGGGCGCGCGGAGTGGCGGTGAGACACGGCGGTGAGCGTGTGGGCTTCGGC 720
QY 721 GTGCTCATCTTCTGGTGTCTCACGGCAACTTCCGTTGGGAGGGCGGTGCGGGCGCC 780
Db 721 GTGCTCATCTTCTGGTGTCTCACGGCAACTTCCGTTGGGAGGGCGGTGCGGGCGCC 780
QY 781 GCCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCCGCTGCCGGGCTCCTTCGCG 840
Db 781 GCCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCCGCTGCCGGGCTCCTTCGCG 840
QY 841 TGGCGCGCTTTCACGAGCGCGCTGCGCATGTTCAGCGCTTACTGGCCCTGAGGCC 900
Db 841 TGGCGCGCTTTCACGAGCGCGCTGCGCATGTTCAGCGCTTACTGGCCCTGAGGCC 900
QY 901 GAGCGCGCGCGCGCTGACCGCGCGGCAAGCCCGCGGGGACCGCCCGCGCGCGCC 1020
Db 901 GAGCGCGCGCGCGCTGACCGCGCGGCAAGCCCGCGGGGACCGCCCGCGCGCGCC 1020
QY 1021 GGGCCACTGCGCCTCGAGGCGCTGGGCTGTTCAGCGGACGCTGCTGACCGGAGCGGC 1080
Db 1021 GGGCCACTGCGCCTCGAGGCGCTGGGCTGTTCAGCGGACGCTGCTGACCGGAGCGGC 1080
QY 1081 AGCGCTCCCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1140
Db 1081 AGCGCTCCCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1140
QY 1141 GTGCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 GTGCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 GCGCGGACCGGCGCGCGCGGACAAAGAGCAAGGCGAGGTGCTGCTGCGCACGCCATC 1260
Db 1201 GCGCGGACCGGCGCGCGCGGACAAAGAGCAAGGCGAGGTGCTGCTGCGCACGCCATC 1260
QY 1261 GAGATCTGCGTC 1272
```

```
Db 1261 GAGATCTGCGTC 1272
RESULT 8
AX455337
LOCUS AX455337 2893 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0210401.
ACCESSION AX455337
VERSION AX455337.1 GI:21714510
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Meyers,R., Kapeller-Libermann,R. and Silos-Santiago,I.
TITLE 18431 and 32374, human protein kinase family members and uses
therefor
JOURNAL Patent: WO 0210401-A 1 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..2893
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
274..1314
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38235.1"
/db_xref="GI:21714511"
/translation="MALKFNKSKTKLKNFLRVSITNSLSSPFFIIKVDVVPETED
CYVFAQYAPAGDLFDIIPQVGLPEDIKRCVQQLGLALDFMHRGLVHRDIKPEV
LLFDRRCRVKLADFGMTRVYGVCRVKVSGTIIPTAPEVCQAGRADGLAVDTGVVWA
FGLVIFCVLTGNFPWEAASGADAFEFVFWQRGLPLGFSQWRFTPEPALRMPQRL
ALPERRGPKEVFRFLKHELTSELRRPSSHARKPGDRPPAAGPLRLRAPGLKRT
VLRESGSRPAPPAVGSVFLPVFPVPVFPVPEPLAPQGPGRITDGRADSKG
QVLIATAIEICV"
BASE COUNT 500 a 989 c 929 g 466 t 9 others
ORIGIN
```

Query Match 88.8%; Score 1131.8; DB 6; Length 2893;

Best Local Similarity 94.8%; Pred. No. 6.7e-135;

Matches 1205; Conservative 0; Mismatches 57; Indels 9; Gaps 3;

QY 14 GCCCAGAGCCTGAGCCGCCCGCTCCCTGACCTGCTGTGGCCGGGACTGCCCTCGGC 73

Db 44 GCCTTGAACCGCCCGCTCCTTGACCTTGTGGCCGAAGGACTGCCCTCGGC 103

QY 74 CTGGTGGC---GTGTGCCCTTCTCACTGAGACATGAGGCCCTGACTCTCGGCAC 130

Db 104 CTGGTGGC---GTGTGCCCTTCTCACTGAGACATGAGGCCCTGACTCTCGGCAC 163

QY 131 TGG---CGCCAGCAGCAGCTCACCAGCACTACGAACTAGT---CCGGAGCTGGGCAAG 184

Db 164 TTGGCCCGCAGCAGCTCACCAGCACTACGAACTAGTCCCGGAGCTGGGCAAG 223

QY 185 GCACCTATGGGAAGTTGACCTGGTGTCTACAAGGCGCAGGACCAAAAATGCACTGA 244

Db 224 GCACCTACGGGAAGTTGATCTGGTGTCTACAAGGCGCAGGACCAAAAATGCACTGA 283

QY 245 AGTTTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGTGAGCATACCA 304

Db 284 AGTTTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGTGAGCATACCA 343

QY 305 ACAGCTCTCTCCAGCCCTTTCATCATCAAGGTCCTTTGACGTGCTTTGAGACAGG 364

Db 344 ACAGCTCTCTCCAGCCCTTTCATCATCAAGGTCCTTTGACGTGCTTTGAGACAGG 403

QY 365 ACTGCTACGCTTTTCCCGAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCTC 424

Db 404 ACTGCTACGCTTTTCCCGAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCTC 463

REFERENCE 1
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase-like protein
JOURNAL Patent: WO 02083882-A 9 24-OCT-2002;
Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1.1982
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 676 c 675 g 308 t
ORIGIN
Query Match 99.9%; Score 1273.4; DB 6; Length 1982;
Best Local Similarity 99.9%; Pred. No. 7.4e-153;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTCTGTGGGCGGGG 60
Db 357 ATGAGCGTGGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTCTGTGGGCGGGG 416
QY 61 ACTGCCCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 120
Db 417 ACTGCCCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 476
QY 121 CTCGCACACTGGCGCGCGCAGCGCTCACCAGCACTACGAAGCTAGTCCGGGAGCTGGC 180
Db 477 CTCGCACACTGGCGCGCGCAGCGCTCACCAGCACTACGAAGCTAGTCCGGGAGCTGGC 536
QY 181 AAGGCACTATGGAAGGTTGACTGGTGTCTTACAGGCAAGGCAAGGCAAGGCAAGGCA 240
Db 537 AAGGCACTATGGAAGGTTGACTGGTGTCTTACAGGCAAGGCAAGGCAAGGCAAGGCA 596
QY 241 CTGAAGTTCTGACACAGACCAACCAAGCTGAGCACTTCTTACGGAGCTGACATC 300
Db 597 CTGAAGTTCTGACACAGACCAACCAAGCTGAGCACTTCTTACGGAGCTGACATC 656
QY 301 ACCAACAGCTCTCTCCAGCGCCCTTATCATCAAGGTCCTTTGACGCTGTTTGAGACA 360
Db 657 ACCAACAGCTCTCTCCAGCGCCCTTATCATCAAGGTCCTTTGACGCTGTTTGAGACA 716
QY 361 GAGACTGCTACGCTTTTCCAGAGTACGACCTGCTGGGACCTGTTGACATCATC 420
Db 717 GAGACTGCTACGCTTTTCCAGAGTACGACCTGCTGGGACCTGTTGACATCATC 776
QY 421 CCTCCAGAGTGGGCTCCCTGAGACACGGTGAAGCTGTGTGACAGCTGGGCTGGGCTG 480
Db 777 CCTCCAGAGTGGGCTCCCTGAGACACGGTGAAGCTGTGTGACAGCTGGGCTGGGCTG 836
QY 481 GCGTGGACTTCATGACGGGCGGAGCTGTGTGACACCGGACATCAAGCCGAGAACCTG 540
Db 837 GCGTGGACTTCATGACGGGCGGAGCTGTGTGACACCGGACATCAAGCCGAGAACCTG 896
QY 541 CTGCTGTTGACCGGAGTGGCGCGCTTAAGCTGGCGGCTTGGGCTGACGATGACGCGCGC 600
Db 897 CTGCTGTTGACCGGAGTGGCGCGCTTAAGCTGGCGGCTTGGGCTGACGATGACGCGCGC 956
QY 601 CTGGGCTCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGGCGCTGAGGTGTGC 660
Db 957 GTGGGCTCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGGCGCTGAGGTGTGC 1016
QY 661 CAGGCGGCGCGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 720
Db 1017 CAGGCGGCGCGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1076
QY 721 GTGCTCATTTCTGGTGTCTACCGGCACTTCCGCTGGGAGCGGCTGGGAGCTGGGAG 780
Db 1077 GTGCTCATTTCTGGTGTCTACCGGCACTTCCGCTGGGAGCGGCTGGGAGCTGGGAG 1136
QY 781 GCCTTCTTCGAGAGTTCTGTGGCTGGCAGCGGGGCGGCTGCGGGGGCTGCTTCGAG 840
Db 1137 GCCTTCTTCGAGAGTTCTGTGGCTGGCAGCGGGGCGGCTGCGGGGGCTGCTTCGAG 1196

QY 841 TGGCGCGGCTTCAAGAGCCCGCGCTGGCATGTTTCCAGCGCTTACTGGCCCTTGAGCCCC 900
Db 1197 TGGCGCGGCTTCAAGAGCCCGCGCTGGCATGTTTCCAGCGCTTACTGGCCCTTGAGCCCC 1256
QY 901 GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Db 1257 GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1316
QY 961 CTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db 1317 CTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1376
QY 1021 GGGCCACTGCGGCTCGAGGCGCTGGGCGGCTTCAAGCGGACGGTGTGACCGAGAGCGGC 1080
Db 1377 GGGCCACTGCGGCTCGAGGCGCTGGGCGGCTTCAAGCGGACGGTGTGACCGAGAGCGGC 1436
QY 1081 AGCGGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Db 1437 GGGGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1496
QY 1141 GTGCCAGTCCCGTGGCGGCTGTGTGCGGAGCCCGGCTAGCTCCCGAGGGGCGGCGGCGG 1200
Db 1497 GTGCCAGTCCCGTGGCGGCTGTGTGCGGAGCCCGGCTAGCTCCCGAGGGGCGGCGGCGG 1556
QY 1201 GGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
Db 1557 GGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1616
QY 1261 GAGATCTGCGTCTGA 1275
Db 1617 GAGATCTGCGTCTGA 1631
RESULT 7
AX589472 1272 bp DNA linear PAT 24-JAN-2003
LOCUS Sequence 1 from Patent WO02083882.
DEFINITION AX589472
ACCESSION AX589472
VERSION AX589472.1 GI:27900910
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase-like protein
JOURNAL Patent: WO 02083882-A 1 24-OCT-2002;
Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1.1272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 191 a 450 c 430 g 201 t
ORIGIN
Query Match 99.0%; Score 1262.4; DB 6; Length 1272;
Best Local Similarity 99.5%; Pred. No. 2.1e-151;
Matches 1266; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTCTGTGGGCGGGG 60
Db 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTCTGTGGGCGGGG 60
QY 61 ACTGCCCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 120
Db 61 ACTGCCCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 120
QY 121 CTCGCACACTGGCGCGCGCAGCGCTCACCAGCACTACGAAGCTAGTCCGGGAGCTGGC 180
Db 121 CTCGCACACTGGCGCGCGCAGCGCTCACCAGCACTACGAAGCTAGTCCGGGAGCTGGC 180

QY	661	CAGGCGGCGCGCGCGAGCGGCTGGCGTGGACACGCGCGTGGAGCTGTGGCGCTTCGGC	720	Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	1434	CAGGCGGCGCGCGCGAGCGGCTGGCGTGGACACGCGCGTGGAGCTGTGGCGCTTCGGC	1493	
QY	721	GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC	780	
Db	1494	GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC	1553	
QY	781	GCTTCTTCGAGAGTCTTGGTGGCTGGCAGCGGGGCGGCTTGGCGGGCTGCTTCGCGAG	840	
Db	1554	GCTTCTTCGAGAGTCTTGGTGGCTGGCAGCGGGGCGGCTTGGCGGGCTGCTTCGCGAG	1613	
QY	841	TGGCGCGCTTACCGAGCCGCGCTGGCGCATGTTCCAGCGCTTACTGGCCCTGGAGGCC	900	
Db	1614	TGGCGCGCTTACCGAGCCGCGCTGGCGCATGTTCCAGCGCTTACTGGCCCTGGAGGCC	1673	
QY	901	GAGCGCGCGCGCGCGAGCGAGGTGTCCGCTTCCCTCAGCAGAGCTACGCTCCGAG	960	
Db	1674	GAGCGCGCGCGCGCGAGCGAGGTGTCCGCTTCCCTCAGCAGAGCTACGCTCCGAG	1733	
QY	961	CTGGCGCGCGCGCGCTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG	1020	
Db	1734	CTGGCGCGCGCGCGCTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG	1793	
QY	1021	GGGCGACTGGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGCGCG	1080	
Db	1794	GGGCGACTGGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGCGCG	1853	
QY	1081	AGCGGCTCCCG	1140	
Db	1854	AGCGGCTCCCG	1913	
QY	1141	GTGCGAGTGGCGCTGCG	1200	
Db	1914	GTGCGAGTGGCGCTGCG	1973	
QY	1201	GGCGGACCGAGCG	1260	
Db	1974	GGCGGACCGAGCG	2033	
QY	1261	GAGATCTGGCTGA	1275	
Db	2034	GAGATCTGGCTGA	2048	
RESULT 4				
LOCUS	AX166553	1275 bp	DNA	linear
DEFINITION	Sequence 44 from Patent WO0138503.			PAT 22-JUN-2001
ACCESSION	AX166553			
VERSION	AX166553.1	GI:14546898		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1. Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Flanagan, P. and Clary, D.S.			
TITLE	Novel human protein kinases and protein kinase-like enzymes			
JOURNAL	Patent: WO 0138503-A 44 31-MAY-2001;			
FEATURES	Sugen, Inc. (US)			
source	Location/Qualifiers			
	1..1275			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
BASE COUNT	191 a	450 c	434 g	200 t
ORIGIN				
Query Match	99.9%; Score 1273.4; DB 6; Length 1275;			
Best Local Similarity	99.9%; Pred. No. 8.3e-153;			

241 CTGAGTTTGTACACAGCAGCAGACCAAGCTGAGAGCTTCCTACCGGAGGTGAGCATC 300
|||
357 CTGAAGTTTGTGAACAAGAGCAGACCAAGCTGAGAGCTTCCTACCGGAGGTGAGCATC 416
|||
301 ACCAACAGCCTCTCTCCAGCGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 360
|||
417 ACCAACAGCCTCTCTCCAGCGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 476
|||
361 GAGGACTGCTAGTTTTCAGAGGAGTACGACCTGCTGGGAGCTGTTGACATCATC 420
|||
477 GAGGACTGCTAGTTTTCAGAGGAGTACGACCTGCTGGGAGCTGTTGACATCATC 536
|||
421 CTTCCCGAGGTGGGCTCCCTGAGACACGGTGAAGCTGTGTGACAGCTGGGCTG 480
|||
537 CTTCCCGAGGTGGGCTCCCTGAGACACGGTGAAGCTGTGTGACAGCTGGGCTG 596
|||
481 GCGCTGGACTTCATGACAGGCGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
|||
597 GCGCTGGACTTCATGACAGGCGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 656
|||
541 CTGCTGTTCGACCGGAGTGCCTGAGCAGCTGTAAGCTGAGCTTCGCGATGACGCGCGC 600
|||
657 CTGCTGTTCGACCGGAGTGCCTGAGCAGCTGTAAGCTGAGCTTCGCGATGACGCGCGC 716
|||
601 GTGGCTGCGCGTCAAGCGCTGAGCGGCTGAGCGCACCATCCCTTACACGCGCCTGAGGTGTC 660
|||
717 GTGGCTGCGCGTCAAGCGCTGAGCGGCTGAGCGCACCATCCCTTACACGCGCCTGAGGTGTC 776
|||
661 CAGCGCGCGCGCGGCTGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 720
|||
777 CAGCGCGCGCGCGGCTGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 836
|||
721 GTGCTCATCTTCTGCTGTCTACCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 780
|||
837 GTGCTCATCTTCTGCTGTCTACCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 896
|||
781 GCTTCTTCGAGAGTTCTGCTGTCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 840
|||
897 GCTTCTTCGAGAGTTCTGCTGTCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 956
|||
841 TGGCGCGGCTTTCAGCGAGCGCGCTGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 900
|||
957 TGGCGCGGCTTTCAGCGAGCGCGCTGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1016
|||
901 GAGCGCGGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 960
|||
1017 GAGCGCGGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1076
|||
961 CTGCGCGGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1020
|||
1077 CTGCGCGGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1136
|||
1021 GGGCCACTGCGCTCGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1080
|||
1137 GGGCCACTGCGCTCGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1196
|||
1081 ACGGCTCCCGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1140
|||
1197 ACGGCTCCCGCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1256
|||
1141 GTGCCAGTCCCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1200
|||
1257 GTGCCAGTCCCGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1316
|||
1201 GGGCCAGCAGCGGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1260
|||
1317 GGGCCAGCAGCGGCGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGCGC 1376
|||
1261 GAGATCTGCTGTA 1275
|||
1377 GAGATCTGCTGTA 1391
|||

RESULT 3
AX364846
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2598
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 446 a 952 c 862 g 338 t
ORIGIN
Query Match 100.0%; Score 1275; DB 6; Length 2598;
Best Local Similarity 100.0%; Pred. No. 4.4e-153;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCCCTCCCTGACCTGCTGTGGGCGGG 60
Db 774 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCCCTCCCTGACCTGCTGTGGGCGGG 833
QY 61 ACTGCCCTGGGCTGTGCGGTGTCGCCCTTCTCACTGAAGACATCAGGCCCTGACT 120
Db 834 ACTGCCCTGGGCTGTGCGGTGTCGCCCTTCTCACTGAAGACATCAGGCCCTGACT 893
QY 121 CTCGCACACTGGCGCGCAGGAGCTGACCGAGCTGACGAGCTAGTCTGGGAGTGGC 180
Db 894 CTCGCACACTGGCGCGCAGGAGCTGACCGAGCTAGTCTGGGAGTGGC 953
QY 181 AAAGCAGCTTATGGAAGTGTGACCTGTGTGTCTCAAGGCGACAGGACAAATGCA 240
Db 954 AAAGCAGCTTATGGAAGTGTGACCTGTGTGTCTCAAGGCGACAGGACAAATGCA 1013
QY 241 CTGAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACCGGAGTGAAGTC 300
Db 1014 CTGAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACCGGAGTGAAGTC 1073
QY 301 ACCAACAGCCTCTCTCCAGCGCCCTTCATCATCAGGTCTTTGAGTGTCTTTGAGACA 360
Db 1074 ACCAACAGCCTCTCTCCAGCGCCCTTCATCATCAGGTCTTTGAGTGTCTTTGAGACA 1133
QY 361 GAGGACTGCTAGTCTTTGCCAGGAGTACGACCTGCTGGGAGCTCTTTGACATCATC 420
Db 1134 GAGGACTGCTAGTCTTTGCCAGGAGTACGACCTGCTGGGAGCTCTTTGACATCATC 1193
QY 421 CTTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCTG 480
Db 1194 CTTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCTG 1253
QY 481 GGGCTGGACTTCATGACAGGCGGAGCTGTGTGACCGGAGCATCAAGCCCGAGAACGTG 540
Db 1254 GGGCTGGACTTCATGACAGGCGGAGCTGTGTGACCGGAGCATCAAGCCCGAGAACGTG 1313
QY 541 CTGCTGTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGCTTCCGATGACGCGCGC 600
Db 1314 CTGCTGTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGCTTCCGATGACGCGCGC 1373
QY 601 GTGGGCTGCCGCTCAAGCGGCTGAGCGGCGACCATCCCTTACACGCGCCTGAGGTGTC 660
Db 1374 GTGGGCTGCCGCTCAAGCGGCTGAGCGGCGACCATCCCTTACACGCGCCTGAGGTGTC 1433
|||

```
Lexicon Genetics Incorporated (US)
Location/Qualifiers
1. .1275
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 192 a 450 c 433 g 200 t
ORIGIN
Query Match 100.0%; Score 1275; DB 6; Length 1275;
Best Local Similarity 100.0%; Pred. No. 5,2e-153;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGTGCCAGAGCTGAGCGCGCCGCTCCCTGACTGCTGTGGGCGGGG 60
DB 1 ATGAGCGTGGGTGCCAGAGCTGAGCGCGCCGCTCCCTGACTGCTGTGGGCGGGG 60
QY 61 ACTGCCCTGGGCTGTGCGGTGTGCCCTTCTCACTGAGACATGAGGCTGACT 120
DB 61 ACTGCCCTGGGCTGTGCGGTGTGCCCTTCTCACTGAGACATGAGGCTGACT 120
QY 121 CTCGCCACACTGGCGCGCAGAGCTGACCAAGCACTACGAAGTGTGCGGAGCTGGG 180
DB 121 CTCGCCACACTGGCGCGCAGAGCTGACCAAGCACTACGAAGTGTGCGGAGCTGGG 180
QY 181 AAAGGCACCTATGGGAAGTTGACCTGTGTGCTTACAAAGGGCACAGGCAAAAATGGCA 240
DB 181 AAAGGCACCTATGGGAAGTTGACCTGTGTGCTTACAAAGGGCACAGGCAAAAATGGCA 240
QY 241 CTGAAGTTGTGAACAAGGCAAAACCAAGCTGGAAGACTTCTACGGGAGTGAGCATC 300
DB 241 CTGAAGTTGTGAACAAGGCAAAACCAAGCTGGAAGACTTCTACGGGAGTGAGCATC 300
QY 301 ACCAAGAGCTCTCTCCAGCGCCCTTCATCAAGGTCTTTGAGTGTCTTTGAGACA 360
DB 301 ACCAAGAGCTCTCTCCAGCGCCCTTCATCAAGGTCTTTGAGTGTCTTTGAGACA 360
QY 361 GAGGACTGACTCTTTGCGGAGGAGTACGACCTGTGTTGAGATCATC 420
DB 361 GAGGACTGACTCTTTGCGGAGGAGTACGACCTGTGTTGAGATCATC 420
QY 421 CTTCCCGAGTGGGCTCCCTGAGGACAGGTAAGCGCTGTGACAGAGTGGGCTG 480
DB 421 CTTCCCGAGTGGGCTCCCTGAGGACAGGTAAGCGCTGTGACAGAGTGGGCTG 480
QY 481 GCGCTGGACTTCATGACGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
DB 481 GCGCTGGACTTCATGACGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
QY 541 CTGCTGTTCGACGCGGAGTGGCGCGCTAAAGCTTGGCGGACTTGGGATGACGCGCG 600
DB 541 CTGCTGTTCGACGCGGAGTGGCGCGCTAAAGCTTGGCGGACTTGGGATGACGCGCG 600
QY 601 GTGGGTGCGCGCTCAAGCGCTGAGCGGACCATCTTACAGCGGCTCAGGTGTGC 660
DB 601 GTGGGTGCGCGCTCAAGCGCTGAGCGGACCATCTTACAGCGGCTCAGGTGTGC 660
QY 661 CAGGCGGCGCGCGGAGTGGCGGTGAGACGCGGCTGAGAGTGTGGGCTTCGAG 720
DB 661 CAGGCGGCGCGCGGAGTGGCGGTGAGACGCGGCTGAGAGTGTGGGCTTCGAG 720
QY 721 GTGCTCATCTTCTGCGTGTCTACCGGCAACTTCCCGTGGGAGCGCGCTGCGGCGCGAC 780
DB 721 GTGCTCATCTTCTGCGTGTCTACCGGCAACTTCCCGTGGGAGCGCGCTGCGGCGCGAC 780
QY 781 GCTTCTTCGAGAGTGTGCTGCGTGGACGCGGCGCGCTGCGGCGGTGCTTCGAG 840
DB 781 GCTTCTTCGAGAGTGTGCTGCGTGGACGCGGCGCGCTGCGGCGGTGCTTCGAG 840
QY 841 TGGCGCGGCTTACCGAGCGCGCTGCGCATGTTCACAGGCTTACTGCGCTTGGAGCC 900
DB 841 TGGCGCGGCTTACCGAGCGCGCTGCGCATGTTCACAGGCTTACTGCGCTTGGAGCC 900
```

```
QY 901 GAGCGCGCGCGCCAGCAAGAGAGTGTTCGGCTTCTCAAGCACGAGCTCACTCGAG 960
DB 901 GAGCGCGCGCGCCAGCAAGAGAGTGTTCGGCTTCTCAAGCACGAGCTCACTCGAG 960
QY 961 CTGCGCGCGCGCGCTCGCACCGCGCGCAAGCGCGCGCGGACCGCGCGCGCGCC 1020
DB 961 CTGCGCGCGCGCGCTCGCACCGCGCGCAAGCGCGCGCGGACCGCGCGCGCGCC 1020
QY 1021 GGGCACTGCGCTCGAGGCGCTTGGCGCTCAAGCGGACGCTGTCGAGAGAGCGGC 1080
DB 1021 GGGCACTGCGCTCGAGGCGCTTGGCGCTCAAGCGGACGCTGTCGAGAGAGCGGC 1080
QY 1081 AGGGCTTCGCGCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1140
DB 1081 AGGGCTTCGCGCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1140
QY 1141 GTGCGAGTGCCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1200
DB 1141 GTGCGAGTGCCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1200
QY 1201 GCGCGGACCGACGCGCGCGCGGACAGAGCAAGGCGAGGTGCTGCGGCAAGCGCATC 1260
DB 1201 GCGCGGACCGACGCGCGCGCGGACAGAGCAAGGCGAGGTGCTGCGGCAAGCGCATC 1260
QY 1261 GAGATCTGCGTCTGA 1275
DB 1261 GAGATCTGCGTCTGA 1275
RESULT 2
AX458305
LOCUS AX458305
DEFINITION Sequence 3 from Patent WO0246428.
ACCESSION AX458305
VERSION AX458305.1 GI:21725000
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Walke D.W., Maricar M., Yu X. and Fridele C.J.
TITLE Novel human kinase and polynucleotides encoding the same
JOURNAL Patent: WO 0246428-A 3 13-JUN-2002;
Lexicon Genetics Incorporated (US);
FEATURES
source
1. .1473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 215 a 533 c 509 g 216 t
ORIGIN
Query Match 100.0%; Score 1275; DB 6; Length 1473;
Best Local Similarity 100.0%; Pred. No. 5e-153;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGTGCCAGAGCTGAGCGCGCCGCTCCCTGACTGCTGTGGGCGGGG 60
DB 117 ATGAGCGTGGGTGCCAGAGCTGAGCGCGCCGCTCCCTGACTGCTGTGGGCGGGG 176
QY 61 ACTGCCCTGGGCTGTGCGGTGTGCCCTTCTCACTGAGACATGAGGCTGACT 120
DB 177 ACTGCCCTGGGCTGTGCGGTGTGCCCTTCTCACTGAGACATGAGGCTGACT 236
QY 121 CTCGCCACACTGGCGCGCAGAGCTGACCAAGCACTACGAAGTGTGCGGAGCTGGG 180
DB 237 CTCGCCACACTGGCGCGCAGAGCTGACCAAGCACTACGAAGTGTGCGGAGCTGGG 296
QY 181 AAAGGCACCTATGGGAAGTTGACCTGTGTGCTTACAAAGGGCACAGGCAAAAATGGCA 240
DB 297 AAAGGCACCTATGGGAAGTTGACCTGTGTGCTTACAAAGGGCACAGGCAAAAATGGCA 356
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 14:17:53 ; Search time 3229 Seconds
(without alignments)
16153.538 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atgagcgtgggtgccaga.....ccatgagatcgtgtcga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_mu.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1275	100.0	1275	6	AX458303	Sequence
2	1275	100.0	1473	6	AX458305	Sequence
3	1275	100.0	2598	6	AX364846	Sequence
4	1273.4	99.9	1275	6	AX166553	Sequence
5	1273.4	99.9	1275	6	AX589478	Sequence
6	1273.4	99.9	1982	6	AX589480	Sequence
7	1262.4	99.0	1272	6	AX589472	Sequence
8	1131.8	88.8	2893	6	AX455337	Sequence
9	1041	81.6	1041	6	AX455339	Sequence
10	966.6	75.8	3858	10	BC024114	Mus muscu
11	966.6	75.8	3953	10	BC025837	Mus muscu
12	966.6	75.8	4123	10	BC031759	Mus muscu
13	963.4	75.6	1527	6	E33454	Novel mamma
14	963.4	75.6	1529	10	AB010154	Rattus no
15	931.2	73.0	936	6	AX589475	Sequence
16	923	72.4	246212	2	AC109769	Rattus no
17	851.2	66.8	136331	9	AC026454	Homo sapi
18	851.2	66.8	163794	2	AC138922	Homo sapi
19	851.2	66.8	178662	2	AC137501	Homo sapi
20	851.2	66.8	178662	2	AC138904	Homo sapi
21	851.2	66.8	185165	2	AC134309	Homo sapi
22	851.2	66.8	230510	9	AC084794	Homo sapi
23	849.6	66.5	7301	6	AX364848	Sequence
24	767.6	60.2	788	6	AX317777	Sequence
25	767.6	60.2	788	9	AF484569	Homo sapi
26	620.8	48.7	227559	2	AC135809	Mus muscu
27	609.6	47.8	231198	2	AC126892	Rattus no
28	609.6	47.8	255924	2	AC096518	Rattus no
29	421.4	33.1	2428	5	AF265347	Danio rer
30	416	32.6	110000	2	AC138525	Continuation (2 of
31	246.4	19.3	156720	2	EX000348	Danio rer
32	236.6	18.6	1047	6	AX481605	Sequence
33	236.6	18.6	1247	6	AX481603	Sequence
34	236.6	18.6	1716	6	AX166515	Sequence
35	229	18.0	110000	2	AC138525	Homo sapi
36	208.8	16.4	220349	2	AC136438	Homo sapi
37	196	15.4	1119	6	AX166516	Sequence
38	179	14.0	1887	5	XLGSPK	Sequence
39	179	14.0	1887	6	AX589476	Sequence
40	162.8	12.8	223879	9	AC008735	Homo sapi
41	158.8	12.5	156506	2	AC068038	Homo sapi
42	149.4	11.7	101451	2	AC138105	Mus muscu
43	147.2	11.5	1992	3	AY071597	Drosophil
44	147.2	11.5	76854	3	AC003052	Drosophil
45	147.2	11.5	86398	2	AC017785	Drosophil

ALIGNMENTS

RESULT 1	AX458303	1275 bp	DNA	linear	PAT 08-JUL-2002
LOCUS	AX458303	Sequence 1 from Patent WO0246428.			
DEFINITION	AX458303	Sequence 1 from Patent WO0246428.			
ACCESSION	AX458303	GI:21724999			
VERSION	AX458303.1				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Walke,D.W., Maricar,M., Yu,X. and Fiddle,C.J.				
TITLE	Novel human kinase and polynucleotides encoding the same				
JOURNAL	Patent: WO 0246428-A 1 13-JUN-2002;				

595	TTCTGTACAGTTCTGTGGGCTCTCGTGGGGCGCACATGACCATCGTGCAGGCC	654
340	TTTACAGTGTCTTTGAGACAGAGGACTGTAGTCTTTTCCAGGAGTAGTACGCACCTGCT	399
655	TACGGCATTTGGCATCGAGTCGGCACACTCTACAGCTTCTGTACGAGAGCCGCTCTGCAC	714
400	GGGGACCTGTTGACATCATCCTCCCAAGTGGGCTCCCTGAGGACACGCTGGAAGCCG	459
715	GGGGACCTCATGGCTTTCATCAGCCCAAGGTGGGCTCCCGAGCCCGGGTGCACCCG	774
460	TGTGTGACAGAGTGGGCTTGGGCTTGAGCTTCATGCACGGGGCGGAGCTGSGTGCACCCG	519
775	TGCGCCGCCAGCTGGCTCCCGCCTTGGAGTATACATCCACGCCCGGSCCTGGTGTACCGG	834
520	GACATCAAGCCCGAGACGTGCTCTGTTTTCGACCGGAGTGGCCGCGCTAAAGCTTGCC	579
835	GACCTGAAGCCGGAGAACGTCTGGTGTGCAGCCCGGCTTCGCGGGGCTTTCAGAGTGACC	894
580	GACTTCGGCATGACGCGCCCGTGGGCTGGCGGTCAAGCGCGTGAAGCGACCATCCCT	639
895	GACTTCGGCCACACAGAGGCTTCGGGGAGCGCTGTGCGCTTGCGCGGCGCCCATCCCC	954
640	TACACGGCGCTGAGTGTGC---CAGGGGGCGCCCGACAGGGCTGGGGGTGGACAG	696
955	TACACGGCCCCGAGCTCTGCGGCCCGCCCGCTCCCGAGGCTTCGCCATTCAGCCC	1014
697	GGCGTGACAGTGTGGGCTTTCGGCGTGTCTCATCTTCTGCGTGTCTACCGGCAACTCCCG	756
1015	GCCCTGACGCTGGCGCTGGGGTCTGTCTTCTGCTCTCTACAGGCTACTTCCCC	1074
757	TGGGA---GGCGGCTGGGGCCGAGCGCTTCTTCGAGGAGTTCGTGCGCTGGCAGCGG	813
1075	TGGACCGGGCCCTGGCGGAGGCGAGCCCTTCTACGAGAGCTTCCTCATCTGGCAGGCG	1134
814	GGCGCCTTCGCGGGCTGCCTTTCGAGTGGCGCGCTTCACCGAGCCCGGCTGGCGCATG	873
1135	TCGGGCGAGCCCGGAGACCGCCCTCAGCCCTGGTTGGCTTGCGCCCGCGGCGGACGCG	1194
874	TTCCAGGCTTACTG	888
1195	CTTCTGGGGGCTG	1209

Search completed: December 7, 2003, 15:17:47
Job time : 303 secs

CC e.g. inflammation, hyperalgesia. The invention is also useful in
CC chromosome mapping, tissue typing and in forensic identification of a
CC biological sample. 58848 molecules are useful as markers of disorders or
CC disease states, as markers of drug activity, or as markers of the
CC pharmacogenomic profile of the subject. The present nucleic acid sequence
CC encodes the human novel protein kinase 58848 protein of the invention.

XX
SQ Sequence 1247 BP; 175 A; 454 C; 421 G; 197 T; 0 other;

Query Match 18.6%; Score 236.6; DB 24; Length 1247;
Best Local Similarity 57.9%; Pred. No. 5.5e-31;
Matches 460; Conservative 0; Mismatches 329; Indels 6; Gaps 2;

QY 100 GAAGACATGAGGCGCTGACTCTCCGACACTGCGCCAGCGACGTCACCAAGCACTAC 159
DB |||||
QY 170 GAGGACATGATGACCTGAGTGTCTACAGCCCTGTCGAGCCGAGGTGGAGGCTCTAC 229
DB |||||
QY 160 GAACCTAGTCCGGAGCTGGGCAAGGACCTATGGAAGGTTGACCTGGTGGTCTACAAG 219
DB |||||
QY 230 GAGGAAGTGGTCCCTGGGCGAGGTCGCTATGCGCGGCTCTCTGTTGTCACCCATCGT 289
DB |||||
QY 220 GGCACAGGCACAAATAGGCACTGAAGTTGTGACAGAGGCAAAACCAAGCTGAAGAC 279
DB |||||
QY 290 CAGAAAGGCACACCCCTGGCACTGAAGAGCTCCCGAAACCCCGCACTGCCCTCGTGGC 349
DB |||||
QY 280 TTCCTACGGAGGTGAGCATCACCAAGAGCTCTCTCCAGCCCTTCATCATCAAGGTC 339
DB |||||
QY 350 TTCCTGACGAGTCTGTGTGGGCTCTCGTGGCGCGCACTCAGCCATCGTACGGCC 409
DB |||||
QY 340 TTGACCTGCTTTGAGACAGAGGCTGACTGCTCTTTGCGCCAGGAGTACGACCTGCT 399
DB |||||
QY 410 TACGGCATTTGATCGAGTGGCACTCTTACAGCTTCTTGAGCGAGCCGCTCTCGAC 469
DB |||||
QY 400 GGGGACCTGTTGACATCATCTCCCTCCAGCTGGGCTCCCTGAGGACAGCTGAAGGC 459
DB |||||
QY 470 GGGGACCTCATGAGCTTCTATCCAGCCCAAGTGGGCTCTCCGAGCCCGCGTGACCGC 529
DB |||||
QY 460 TGTGTGACGAGCTGGGCTGGCGCTGAGCTTCTATGACGGGCGGCGAGCTGGTCAACGC 519
DB |||||
QY 530 TGGCGCGCGAGTGGCTTCGCGCTCGAGTACATCCACGCGCGGCTGGTGTACCGG 589
DB |||||
QY 520 GACATCAAGCCCGAAGAGCTGCTGTTGACCGCGAGTGGCGCGGCTAAAGTGGCC 579
DB |||||
QY 590 GACCTGAAGCCGGAAGAGCTGCTGTGTGCGACCCCGGCTGCGGCGCTTCAAGCTGACC 649
DB |||||
QY 580 GACTTCGACATGACGCGCGCTGGGCTGCGCTGCGCTCAAGCGGTGAGCGGCACTACCT 639
DB |||||
QY 650 GACTTCGCGCCACACGAGGCTCGCGGAGCTGCTGCGCTGCGCGCGCGGCTTCAAGCTGACC 709
DB |||||
QY 640 TACAGCGGCTGAGGTGTC---CAGGCGGCGCGCGCGCGCTGCGGCTGGGACACG 696
DB |||||
QY 710 TACAGCGGCGCGAGCTCTGCGCGCGCGCGCGCTCCCGAGGCGCTGCCATTCAGGCC 769
DB |||||
QY 697 GCGTGGAGCTGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
DB |||||
QY 770 GCGCTGGAGCGCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
DB |||||
QY 757 TGGGA---GGCGGCTCGGCGCGCGCGCTTCTGAGGAGTGTGCTGCTGGCAGCGG 813
DB |||||
QY 830 TGGGACCGGCGCGCTGGCGGAGCGGCGGCTTCTACGAGACTTCTATCTGCGAGGCG 889
DB |||||
QY 814 GCGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
DB |||||
QY 890 TCGGGCCAGCCCGGAGCCGCTCAGCCCTGCTGCGGCTGCTGCGCGCGCGCGCGCGCGG 949
DB |||||
QY 874 TTCAGGCTTACTG 888
DB |||||
QY 950 CTTCTGCGGGGCTG 964
DB |||||

RESULT 15
AAS06706
ID AAS06706 standard; cdna; 1716 BP.

XX
AC AAS06706;
XX 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #6.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX WO200138503-A2.
XX 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US32085.
XX 24-NOV-1999; 99US-0167482.
XX (SUGE-) SUGEN INC.
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI; 2001-343950/36.
XX P-PSDB; AAU03506.
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX Example 1; Figure 1; 433pp; English.
XX AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (TK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
XX modulators of protein kinase expression and activity.
XX SQ Sequence 1716 BP; 285 A; 597 C; 564 G; 270 T; 0 other;

Query Match 18.6%; Score 236.6; DB 22; Length 1716;
Best Local Similarity 57.9%; Pred. No. 5.5e-31;
Matches 460; Conservative 0; Mismatches 329; Indels 6; Gaps 2;

QY 100 GAAGACATGAGGCGCTGACTCTCCGACACTGCGCCAGCGACGTCACCAAGCACTAC 159
DB |||||
QY 160 GAACCTAGTCCGGAGCTGGGCAAGGACCTATGGAAGGTTGACCTGGTGGTCTACAAG 219
DB |||||
QY 475 GAGGAAGTGGTCCCTGGGCGAGGTCGCTATGCGCGGCTCTCTGCTGCTCACCATCGT 534
DB |||||
QY 220 GGCACAGGCACAAATAGGCACTGAAGTTGTGACAGAGGCAAAACCAAGCTGAAGAC 279
DB |||||
QY 535 CAGAAAGGCACACCCCTGGCACTGAAGAGCTCCCGAAACCCCGCACTGCCCTCGTGGC 594
DB |||||
QY 280 TTCCTACGGAGGTGAGCATCACCAAGGCTCTCTCCAGCGCCCTTCATCATCAAGGTC 339
DB |||||

gene whose imprinting is lost in cancer, or if the test substance inhibits imprinting of a gene whose imprinting is gained in cancer. The methylated CpG islands are useful for providing an assessment of the risk of developing cancer, or for providing diagnostic information relative to cancer which involves determining the methylation status of the CpG island in a patient's DNA. The EG cells allow the accession of imprinted genes which are useful for detecting birth defects, diabetes and cancers associated with aberrant imprinting. The EG cell lines represent the first in vitro model system in which genomic imprinting can be followed dynamically and the two alleles can be distinguished. AAS21030-AAS21083 represent isolated CpG islands which are not present in public databases.

XX SQ Sequence 788 BP; 111 A; 291 C; 300 G; 86 T; 0 other;

Query Match 60.2%; Score 767.6; DB 24; Length 788;

Best Local Similarity 99.4%; Pred. No. 8e-120;

Matches 781; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 420 CCCTCCCCAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 479

DB 785 CGGCCCGAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 726

QY 480 GCGCTGGAATTCATGCAACGGCGGCGAGCTGTGCACCGGACATCAAGCCCGAGAACT 539

DB 725 GCGCTGGAATTCATGCAACGGCGGCGAGCTGTGCACCGGACATCAAGCCCGAGAACT 666

QY 540 GCTGCTGTTCACCGGAGTGCAGCGGCGGCGGTAAGCTGCGGCTTGGCATGACGCGCG 599

DB 665 GCTGCTGTTCACCGGAGTGCAGCGGCGGCGGTAAGCTGCGGCTTGGCATGACGCGCG 606

QY 600 CGTGGGCTGCGCGCTCAAGCGGCTGAGCGGACCATTCCTTACACGGCGCTGAGGCTGT 659

DB 605 CGTGGGCTGCGCGCTCAAGCGGCTGAGCGGACCATTCCTTACACGGCGCTGAGGCTGT 546

QY 660 CCAGCGCGGCGCGCGGCTGCGGCTGCGGACACGGCGGCTGCGGCTTCCG 719

DB 545 CCAGCGCGGCGCGCGGCTGCGGCTGCGGACACGGCGGCTGCGGCTTCCG 486

QY 720 CGTGTCTATCTTCGCTGCTCACCGGCACTTCCTGCGGAGCGGCTGCGGCGCGGA 779

DB 485 CGTGTCTATCTTCGCTGCTCACCGGCACTTCCTGCGGAGCGGCTGCGGCGCGGA 427

QY 780 CGCCTTCTTCGAGGAGTTCGCTGCGTGCAGCGGCGGCTGCGGCGGCTTCCGCA 839

DB 426 CGCCTTCTTCGAGGAGTTCGCTGCGTGCAGCGGCGGCTGCGGCGGCTTCCGCA 367

QY 840 GTGGCGCGCTTCCAGGCGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 899

DB 366 GTGGCGCGCTTCCAGGCGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 307

QY 900 CGAGCGCGCGGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 959

DB 306 CGAGCGCGCGGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 247

QY 960 GCTGGCGCGGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 1019

DB 246 GCTGGCGCGGCGGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGCC 187

QY 1020 CGGGCCACTGCGCTTCAGGCGGCTTGGCGGCTCAAGCGGACGCTGCTACCGAGAGCG 1079

DB 186 CGGGCCACTGCGCTTCAGGCGGCTTGGCGGCTCAAGCGGACGCTGCTACCGAGAGCG 127

QY 1080 CAGCGGCTCCGCGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1139

DB 126 CAGCGGCTCCGCGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 67

QY 1140 GGTGCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1199

DB 66 GGTGCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 7

QY 1200 CGGCG 1205

DB 6 CGGCG 1

RESULT 14

ABK89200

ID ABK89200 standard; DNA; 1247 BP.

XX AC ABK89200;

DT 21-OCT-2002 (first entry)

XX DE DNA encoding human novel protein kinase 58848 protein.

XX KW Human; protein kinase 58848; protein kinase associated disorder;
KW 58848-associated disorder; cellular disorder; proliferative disorder;
KW differentiate disorder; cancer; leukaemia; hormonal disorder; diabetes;
KW immune disorder; autoimmune disease; blood vessel disorder; hypertension;
KW atherosclerosis; platelet disorder; cardiovascular disorder; ischaemia;
KW cardiac hypertrophy; heart failure; neurological disorder; rickets;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; AIDS;
KW acquired immunodeficiency syndrome; bone metabolism disorder; pain;
KW osteoporosis; cirrhosis; haematopoietic neoplastic disorder; Hepatitis B;
KW Hodgkin's disease; acute leukaemia; liver disorders; Gaucher's disease;
KW viral disease; metabolic disorder; inflammation; hyperalgesia;
KW chromosome mapping; tissue typing; forensic identification;
KW pharmacogenomic profile; gene; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers

5'UTR 1..43

FT CDS 44..1090

FT 3'UTR 1091..1247

FT /product= "Human novel protein kinase 58848 protein"

FT /note= "Specifically claimed in claim 1"

FT /tag= c

XX WO200255713-A2.

XX 18-JUL-2002.

XX 26-NOV-2001; 2001WO-US44346.

XX 08-DEC-2000; 2000US-254401P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller-Libermann R, Acton S;

XX WPI; 2002-590676/63.

XX P-PSDB; ABG30935.

XX Novel human protein kinase polypeptide, designated 58848, useful for treating diseases including cellular, proliferative, bone metabolism, cardiovascular, neurological, and haematopoietic neoplastic disorders

XX Claim 1; Page 97-99; 104pp; English.

XX The present invention relates to a new human protein kinase polypeptide designated 58848. The invention is useful for modulating 58848-mediated activities which are useful for developing diagnostic and therapeutic agents for protein kinase associated or other 58848-associated disorders such as cellular, proliferative and/or differentiate disorders e.g. cancer, leukaemia; hormonal disorders e.g. diabetes; immune disorders e.g. autoimmune disease; blood vessel disorders e.g. atherosclerosis, hypertension; platelet disorders; cardiovascular disorders e.g. cardiac hypertrophy, heart failure; neurological disorders e.g. ischaemia, Alzheimer's disease, Parkinson's disease, Huntington's disease, acquired immunodeficiency syndrome (AIDS); bone metabolism disorders e.g. rickets, osteoporosis, cirrhosis; haematopoietic neoplastic disorders e.g. Hodgkin's disease, acute leukaemia; liver disorders e.g. Gaucher's disease, viral diseases e.g. Hepatitis B; pain or metabolic disorder

DR WPI; 2002-083100/11.

XX Forming embryonic germ cells useful as model system to study imprinting

PT Involves mating genetically divergent male and female mammal of same

PT species, dissecting and dissociating embryo obtained from pregnant

PT mammal -

XX Disclosure; Page 115; 125pp; English.

XX The present invention relates to a model system for genomic imprinting

CC using pluripotent mouse embryonic germ (EG) cell lines derived from an

CC interspecific cross. Also disclosed is a library containing methylated

CC CpG islands and a method for assaying methylation in one or more

CC imprintable genes. The gene imprinting assay is carried out by

CC single-strand conformation polymorphism (SSCP), quantitative sequencing,

CC carried out to determine the post-translational modification of histones.

CC The method further involves identifying a test substance as a candidate

CC drug for treating cancer if the test substance enhances imprinting of a

CC gene whose imprinting is lost in cancer, or if the test substance

CC inhibits imprinting of a gene whose imprinting is gained in cancer. The

CC methylated CpG islands are useful for providing an assessment of the risk

CC of developing cancer, or for providing diagnostic information relative to

CC cancer which involves determining the methylation status of the CpG

CC island in a patient's DNA. The EG cells allow the accession of imprinted

CC genes which are useful for detecting birth defects, diabetes and cancers

CC associated with aberrant imprinting. The EG cell lines represent the

CC first in vitro model system in which genomic imprinting can be followed

CC dynamically and the two alleles can be distinguished. AAS20970-AAS21029

CC represent DNA sequences of unknown function given in the present

CC invention.

CC Note: The present sequence is given in the Sequence listing but is not

CC mentioned elsewhere in the patent.

XX

SQ Sequence 788 BP; 111 A; 291 C; 300 G; 86 T; 0 other;

Query Match 60.2%; Score 767.6; DB 24; Length 788;

Best Local Similarity 99.4%; Pred. No. 9e-120;

Matches 781; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 420 CCTCCCGAGTGGGGCTCCCTGAGGACAGCGTGTAAGCGCTGTGTGACAGCTGGGGCT 479

DB 785 CGCCCGCAGGTGGGGCTCCCTGAGGACAGCGTGTAAGCGCTGTGTGACAGCTGGGGCT 726

QY 480 GCGCTGGACTTCATGCACGCGCGGCGAGCTGTGCACCGCGACATCAAGCCCGAGAACT 539

DB 725 GCGCTGGACTTCATGCACGCGCGGCGAGCTGTGCACCGCGACATCAAGCCCGAGAACT 666

QY 540 GCTGCTGTTCGACCGGAGTGGCGCGGCTAAAGCTGGCGGACTTCGGCATGACGCGCG 599

DB 665 GCTGCTGTTCGACCGGAGTGGCGCGGCTAAAGCTGGCGGACTTCGGCATGACGCGCG 606

QY 600 CTTGGGCTGCGCGTCACGCGGTGAGCGGACCATCCCTTACACGGCGCTGAGGTGTG 659

DB 605 CTTGGGCTGCGCGTCACGCGGTGAGCGGACCATCCCTTACACGGCGCTGAGGTGTG 546

QY 660 CCAGCGGCGCGCGCGGCTGGCGGTGACACGCGGCTGAGCGCTGTGGGCTTCGG 719

DB 545 CCAGCGGCGCGCGCGGCTGGCGGTGACACGCGGCTGAGCGCTGTGGGCTTCGG 486

QY 720 CTTGCTCATCTTCTGGTGTCTACCGGCACTTCCGTTGGGAGCGGCTGGGCGCGCA 779

DB 485 CTTGCTCATCTTCTGGTGTCTACCGGCACTTCCGTTGGGAGCGGCTGGGCGCGCA 427

QY 780 CCGCTTCTTCGAGGAGTTCGTGCTGGCAGCGGCGCGCTTCGCGGGGCTCCCTTCGCA 839

DB 426 CCGCTTCTTCGAGGAGTTCGTGCTGGCAGCGGCGCGCTTCGCGGGGCTCCCTTCGCA 367

QY 840 GTGGGCGCGCTTCACGAGCCCGCGCTGGCGATGTTCCAGCGCTTACTGGCCCTCGAGCC 899

DB 366 GTGGGCGCGCTTCACGAGCCCGCGCTGGCGATGTTCCAGCGCTTACTGGCCCTCGAGCC 307

QY 900 CGAGCGCGCGCGCGCGCAAGAGAGTGTTCGCTTCCTCAAGCAGGACTCACGTCCGA 959

DB 306 CGAGCGCGCGCGCGCGCAAGAGGTGTTCGCTTCCTCAAGCAGGACTCACGTCCGA 247

QY 960 GCTGCGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 1019

DB 246 GCTGCGCGCGCGCGCGCTTCGACCGCGCGCGCGCAAGCGCGCGCGCGCGCGCGCG 187

QY 1020 CGGCGCACTGCGCTCGAGCGCGCTGGCGCTCAAGCGGACCGGTGTGACCGAGAGCGG 1079

DB 186 CGGCGCACTGCGCTCGAGCGCGCTGGCGCTCAAGCGGACCGGTGTGACCGAGAGCGG 127

QY 1080 CAGCGGCTCCCG 1139

DB 126 CAGCGGCTCCCG 67

QY 1140 GGTGCCAGTCCCGCGCGCGCGCTGTGCGCGGCGCGCGCTAGCTCCCGAGGCGCGCC 1199

DB 66 GGTGCCAGTCCCGCGCGCGCGCTGTGCGCGGCGCGCGCTAGCTCCCGAGGCGCGCC 7

QY 1200 CGGCGCG 1205

DB 6 CGGCGCG 1

RESULT 13

AAS21038/c

ID AAS21038 standard; DNA; 788 BP.

XX AAS21038;

XX AC

XX 09-APR-2002 (first entry)

XX DE

XX Isolated CpG island 2-42.

XX Human; genomic imprinting; pluripotent mouse embryonic germ cell line;

XX EG; methylated CpG island; DNA methylation; gene imprinting;

XX post-translational modification of histone; cancer; birth defect;

XX diabetes; aberrant imprinting; ds.

XX Homo sapiens.

XX WO200190313-A2.

XX 29-NOV-2001.

XX 22-MAY-2001; 2001WO-US16253.

XX 22-MAY-2000; 2000US-206158P.

XX 22-MAY-2000; 2000US-206161P.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Feinberg A, Strichman-Almashanu L, Jiang S;

XX WPI; 2002-083100/11.

XX Forming embryonic germ cells useful as model system to study imprinting

PT involves mating genetically divergent male and female mammal of same

PT species, dissecting and dissociating embryo obtained from pregnant

PT mammal -

XX Disclosure; Fig 14; 125pp; English.

XX The present invention relates to a model system for genomic imprinting

CC using pluripotent mouse embryonic germ (EG) cell lines derived from an

CC interspecific cross. Also disclosed is a library containing methylated

CC CpG islands and a method for assaying methylation in one or more

CC imprintable genes. The gene imprinting assay is carried out by

CC single-strand conformation polymorphism (SSCP), quantitative sequencing,

CC single nucleotide primer extension or hot stop PCR. The assays are

CC carried out to determine the post-translational modification of histones.

CC The method further involves identifying a test substance as a candidate

CC drug for treating cancer if the test substance enhances imprinting of a

PR 17-MAY-2000; 2000US-205228P.
PR 12-DEC-2000; 2000US-0734032.
PR 26-MAR-2001; 2001US-0816094.
XX
PA (APPL-) APPLERA CORP.
XX
PI Wei M, Chandramouliswara I, Ye J, Ketchum KA, Di Francesco V;
PI Beasley EM;
XX
DR P-PSDB; AAU10023.
DR WPI; 2002-089857/12.
XX
XX Human kinase protein and polynucleotides encoding them, useful for
PT identifying modulators of kinase polypeptides and for treating,
PT preventing, and/or diagnosing neurodegenerative diseases and cancer -
XX
PS Claim 4; Figure 3; 65pp; English.
XX
CC This sequence represents a human protein kinase N gene of the invention.
CC The invention comprises nucleotide and protein sequences of an isolated
CC protein which is related to the PKN kinase subfamily. The protein
CC kinase N gene is located on chromosome 16. The protein may have
CC cytosolic and neuroprotective and can be used in gene therapy
CC possibly as a human kinase protein expression or activity modulator.
CC The nucleic acids and polypeptides of the invention may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. The nucleic acids (or vectors
CC containing them) and the kinase may be used to treat disorders
CC associated with decreased expression, either by rectifying mutations or
CC deletions in a patient's genome that affect the activity of the enzyme
CC by expressing inactive proteins or to supplement the patients own
CC production of kinases. Additionally, the nucleic acids may be used to
CC produce the kinase, by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The nucleic acid and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides may also be used as antigens in the
CC production of antibodies against the kinase and in assays to identify
CC modulators of kinase expression and activity. The anti-kinase
CC antibodies and antagonists may also be used to down regulate expression
CC and activity. The anti-kinase antibodies may also be used as diagnostic
CC agents for detecting the presence of kinase polypeptides in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). Disorders that may
CC be prevented, diagnosed and/or treated by the above methods include, for
CC example neurodegenerative diseases, inflammation, arteriosclerosis
CC psoriasis, cancer, papilloma virus infection, Alzheimer's disease
CC and growth disorders.
XX
SQ Sequence 7301 BP; 1466 A; 2159 C; 2281 G; 1395 T; 0 other;

Query Match 66.6%; Score 849.6; DB 24; Length 7301;
Best Local Similarity 99.5%; Pred. No. 1.6e-133;
Matches 852; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 420 CCTCTCCAGTGGGCTCCCTGAGACACACGTTGAGCGCTGTCTCAGCAGCTGGGCT 479
DB 3448 CGCGCCGACGAGTGGGCTCCCTGAGGACACGTTGAGCGCTGTCTCAGCAGCTGGGCT 3507
QY 480 GCGCTGGACTTCATGACGGGGCGGAGCTGGTGGACCGGACATCAAGCCCGAGAACGT 539
DB 3508 GCGCTGGACTTCATGACGGGGCGGAGCTGGTGGACCGGACATCAAGCCCGAGAACGT 3567
QY 540 GCTGCTGTTTCGACCGGAGTGGCCCGTAAAGCTTGGCGACTTCGGCATGACGCGCG 599
DB 3568 GCTGCTGTTTCGACCGGAGTGGCCCGTAAAGCTTGGCGACTTCGGCATGACGCGCG 3627
QY 600 CTTGGGCTGCCGCTGAAGCGGTGAGCGGACCATCCCTTACCGGCGCTGAGGTGTG 659
DB 3628 CTTGGGCTGCCGCTGAAGCGGTGAGCGGACCATCCCTTACCGGCGCTGAGGTGTG 3687
QY 660 CCAGGCGGGCGCGCGAGCTGGCGGTGGACACGGGCTGGAGCTGGGCGCTTCGG 719

Db 3688 CCAGGCGGGCGCGCGGACCGGCTGGGCGGTGGACACAGGGGCTGGAGCTGTGGGCTTCGG 3747
QY 720 CGTGTCTCATCTTCTCGCTGTCTCAACGGCAACTTCCGFTGGAGGCGGCTCGGGGCGCGA 779
Db 3748 CGTGTCTCATCTTCTCGCTGTCTCAACGGCAACTTCCGFTGGAGGCGGCTCGGGGCGCGA 3807
QY 780 CGCTTCTTCAGGAGTTCGTGCGCTGACAGGGGCGCGCTGCGGGGCTGCCTTCGCA 839
Db 3808 CGCTTCTTCAGGAGTTCGTGCGCTGACAGGGGCGCGCTGCGGGGCTGCCTTCGCA 3867
QY 840 GTGGCGCGCTTTCACCGAGCGCGGCTCGCATGTTCCAGCGCTTACTGGGCGCTGGAGCC 899
Db 3868 GTGGCGCGCTTTCACCGAGCGCGGCTCGCATGTTCCAGCGCTTACTGGGCGCTGGAGCC 3927
QY 900 CGAGCGCGCGCGCGCGAGCAAGGAGTTCGCGTTCCTAAGCACGAGCTCACGTCGCA 959
Db 3928 CGAGCGCGCGCGCGCGAGCAAGGAGTTCGCGTTCCTAAGCACGAGCTCACGTCGCA 3987
QY 960 GCTGGCGCGCGCGCGCTCGCACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 1019
Db 3988 GCTGGCGCGCGCGCGCGCTCGCACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 4047
QY 1020 CGGGCACTGCGCGCTCGAGGCGCGCTGGGCGCGCTCAAGCGGACGGTGTCTGACGAGCGG 1079
Db 4048 CGGGCACTGCGCGCTCGAGGCGCGCTGGGCGCGCTCAAGCGGACGGTGTCTGACGAGCGG 4107
QY 1080 CAGCGGCTCCG 1139
Db 4108 CAGCGGCTCCG 4167
QY 1140 GGTGCACTGCG 1199
Db 4168 GGTGCACTGCG 4227
QY 1200 CGGGCGGACCGACG 1259
Db 4228 CGGGCGGACCGACG 4287
QY 1260 CGAGATCTGCGTCTGA 1275
Db 4288 CGAGATCTGCGTCTGA 4303
RESULT 12
AAS20990/c
ID AAS20990 standard; DNA; 788 BP.
XX
AC AAS20990;
XX
DT 09-APR-2002 (first entry)
XX
DE DNA sequence #21 relating to gene imprinting invention.
XX
KW Human; genomic imprinting; pluripotent mouse embryonic germ cell line;
KW EG; methylated CpG island; DNA methylation; gene imprinting;
KW post-translational modification of histone; cancer; birth defect;
KW diabetes; aberrant imprinting; ds.
XX
OS Homo sapiens.
XX
PN WO2001903113-A2.
XX
PD 29-NOV-2001.
XX
PF 22-MAY-2001; 2001WO-US16253.
XX
PR 22-MAY-2000; 2000US-206158P.
PR 22-MAY-2000; 2000US-206161P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Feinberg A, Strichman-Almaashanu L, Jiang S;

CC modulating the activity of STPKP in the diseases. These diseases include
CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
CC human STPKP DNA.

XX SQ Sequence 936 BP; 155 A; 308 C; 308 G; 165 T; 0 other;
Query Match 73.0%; Score 931.2; DB 25; Length 936;
Best Local Similarity 99.7%; Pred. No. 3.4e-147;
Matches 933; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGAGCGTGGGTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCGCGGG 60
Dy 1 ATGAGCGTGGGTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCGCGGG 60
Qy 61 ACTGCCCTGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
Dy 61 ACTGCCCTGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
Qy 121 CTCGACACTGGCGCGCCGAGCAGCTCACCAGCACTACGAACTAGTCCGGAGCTGGGC 180
Dy 121 CTCGACACTGGCGCGCCGAGCAGCTCACCAGCACTAGTCCGGAGCTGGGC 180
Qy 181 AAAGCAGCTATGGGAAGTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Dy 181 AAAGCAGCTATGGGAAGTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Qy 241 CTGAAGTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGTGAGCATC 300
Dy 241 CTGAAGTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGTGAGCATC 300
Qy 301 ACCAAGCCTCTCTCCAGCGCCCTTCATCAACAGGCTTTGACGTGTCTTGAGACA 360
Dy 301 ACCAAGCCTCTCTCCAGCGCCCTTCATCAACAGGCTTTGACGTGTCTTGAGACA 360
Qy 361 GAGGAGCTGACGTCTTTGCGCAGAGTACGACCTGCTGGGAGCTGTTGACATATC 420
Dy 361 GAGGAGCTGACGTCTTTGCGCAGAGTACGACCTGCTGGGAGCTGTTGACATATC 420
Qy 421 CTTCCCGAGTGGGTCTCCTGAGACACGCTGAAGCGTGTGTGAGCAGCTGGCGCTG 480
Dy 421 CTTCCCGAGTGGGTCTCCTGAGACACGCTGAAGCGTGTGTGAGCAGCTGGCGCTG 480
Qy 481 GCGCTGGAATTCATGACGGCGGCGAGCTGGTGTGCAACCGCAATCAAGCCCGAGAACGTG 540
Dy 481 GCGCTGGAATTCATGACGGCGGCGAGCTGGTGTGCAACCGCAATCAAGCCCGAGAACGTG 540
Qy 541 CTGCTGTTCCGCGGAGTGGCGCGCTAAAGCTGGCGGCTTCGGCATGACGCGCGC 600
Dy 541 CTGCTGTTCCGCGGAGTGGCGCGCTAAAGCTGGCGGCTTCGGCATGACGCGCGC 600
Qy 601 GTGGGTGCGCGGTCAAGCGCGTGAAGCGGCATCCCTTACAGCGCGCTGAGGTGTC 660
Dy 601 GTGGGTGCGCGGTCAAGCGCGTGAAGCGGCATCCCTTACAGCGCGCTGAGGTGTC 660
Qy 661 CAGCGCGCGCGCGCGCGCGTGGCGGTGGACACCGGGGTGAACGCTGTGGCGCTTCGCG 720
Dy 661 CAGCGCGCGCGCGCGCGCGTGGCGGTGGACACCGGGGTGGACGCTGTGGCGCTTCGCG 720
Qy 721 GTGCTCATCTTCGCTGTGCTCAACGGCACTTCCCGTGGAGCGCGCTGGCGCGGAC 780
Dy 721 GTGCTCATCTTCGCTGTGCTCAACGGCACTTCCCGTGGAGCGCGCTGGCGCGGAC 780
Qy 781 GCCTTCTTCGAGGAGTTCGTGCGCTGCGAGCGGGCGCGCTGCGGGGTGCTTCGCGAG 840
Dy 781 GCCTTCTTCGAGGAGTTCGTGCGCTGCGAGCGGGCGCGCTGCGGGGTGCTTCGCGAG 840
Qy 841 TGGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGGCC 900
Dy 841 TGGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGGCC 900
Qy 901 GAGCGCGCGCGCGCGCGAGGAGTGTTCGCGTTC 936

Db 901 GAGCGCGCGCGCGCGAGGAGTGTTCGCGTTC 936

RESULT 11

AAS17863
ID AAS17863 standard; DNA; 7301 BP.
XX AAS17863;
AC
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein kinase N genomic sequence.
XX
KW Human; ds; protein kinase N; cytostatic; neuroprotective; cancer;
KW gene therapy; antigen; antibody; neurodegenerative disease;
KW inflammation; arteriosclerosis; psoriasis; growth disorder;
KW chromosome 16; papilloma virus infection, Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(282,T)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(335,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 769..4303
FT /*tag= c
FT /product= "Protein Kinase N"
FT exon 769..994
FT /*tag= d
FT /number= 1
FT intron 995..2376
FT /*tag= e
FT /number= 1
FT exon 2377..2579
FT /*tag= f
FT /number= 2
FT intron 2580..3457
FT /*tag= g
FT /number= 2
FT variation replace(2697,A)
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT exon 3458..4300
FT /*tag= i
FT /number= 12
FT variation replace(5406,C)
FT /*tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6309,C)
FT /*tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "This nucleotide can be either C, T or absent"
FT variation replace(6324..6326,CC)
FT /*tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "The nucleotide at position 6325 can be either T or absent"
FT variation replace(6325..6327,TT)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "The nucleotide at position 6326 can be either C or absent"
XX
PN WO200188148-A2.
XX
PD 22-NOV-2001.
XX
PF 17-MAY-2001; 2001WO-US15776.

[illegible]

CC idiopathic hyperglycaemia, haematopoietic disorders such as autoimmune
 CC disease including rheumatoid arthritis, osteoarthritis, multiple
 CC sclerosis, systemic lupus erythematosus, psoriasis, conjunctivitis,
 CC ulcerative colitis, asthma, allergic asthma, autoimmune uveitis,
 CC aplastic anaemia, Grave's disease, graft-versus-host disease; liver
 CC disorders e.g. Gaucher's disease, Wilson's disease; and platelet
 CC disorders. Numerous other diseases that can be treated or prevented
 CC are listed in the specification. The present sequence is the cDNA
 CC encoding protein kinase 32374.
 XX
 SQ Sequence 2893 BP; 500 A; 989 G; 929 G; 466 T; 9 other;
 Query Match 98.8%; Score 1131.8; DB 24; Length 2893;
 Best Local Similarity 94.8%; Pred. No. 9.5e-181;
 Matches 1205; Conservative 0; Mismatches 57; Indels 9; Gaps 3;
 QY 14 GCCCAGAGCCTGAGCCGCCCGCTCCCTGACCTGCTGTGGCCGGGAGCTGCCCTGGGC 73
 Db 44 GCCTTGAACCGGCCCGGCTCTTGACCTTGTGGCCGAGGACTGCCCTGGGC 103
 QY 74 CTGGTGCCG---GTGTGCCCTTCTCACTGAAGACATGACGGCCCTGACTCTCCGCACAC 130
 Db 104 CTGGTCCGGGTGTGCCCTTCTCACTTGAAGACATGACGGCCCTGACTCTCCGCACAC 163
 QY 131 TGG---CCGCAGAGCATCTACCAAGCACTACGAACCTAGT---CCGGGAGCTGGGCAAG 184
 Db 164 TTGGCCCGCCAGCGACGCTCAACCAAGCACCTACGAAACTAGTCCCGGGAGCTGGGCAAG 223
 QY 185 GCACCTATGGAGGTGACCTGTGTCTACAGGGCAGAGGCACACAAAATGCACTGA 244
 Db 224 GCACCTACGGGAGGTGATCTGTGTCTACAGGGCAGAGGCACACAAAATGCACTGA 283
 QY 245 AGTTTGTGAACAGACAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATCA 304
 Db 284 AGTTTGTGAACAGACAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATCA 343
 QY 305 ACAGCCTCTCTCAGGCCCTTCAATCAATCAAGTCTTTGACGTGGTCTTTGAGACAGAG 364
 Db 344 ACAGCCTCTCTCAGGCCCTTCAATCAATCAAGTCTTTGACGTGGTCTTTGAGACAGAG 403
 QY 365 ACTGTACGCTTTTCCCGAGGAGTACGACCTGTGTGGGACCTTTTGACATCATCCTC 424
 Db 404 ACTGTACGCTTTTCCCGAGGAGTACGACCTGTGTGGGACCTTTTGACATCATCCTC 463
 QY 425 CCCAGTGGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGACAGCTGGGCTGGCGC 484
 Db 464 CCCAGTGGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGACAGCTGGGCTGGCGC 523
 QY 485 TGGACTTATGACGGGGGAGCTGTGTGACACCGGACATCAAGCCCGAGACGTGTGC 544
 Db 524 TGGACTTATGACGGGGGAGCTGTGTGACACCGGACATCAAGCCCGAGACGTGTGTGC 583
 QY 545 TGTTCGACCGGAGTGGCCCGCGCTAAAGCTTGGCGACTTCGGCATGACGGCCCGCTGG 604
 Db 584 TGTTCGACCGGAGTGGCCCGCGCTAAAGCTTGGCGACTTCGGCATGACGGCCCGCTGG 643
 QY 605 GCTCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACACGGCGCTGAGTGTGCCAGG 664
 Db 644 GCTCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACACGGCGCTGAGTGTGCCAGG 703
 QY 665 CGGGCCGCGCAGCGGGTGGCGGTGGACACGGGGTGGAGCTGTGGGCTTCGGCGTGC 724
 Db 704 CGGGCCGCGCAGCGGGTGGCGGTGGACACGGGGTGGAGCTGTGGGCTTCGGCGTGC 763
 QY 725 TCATCTTCTGTGCTCACCGCACTTCCCGTGGGAGCGCGTGGCGCGCGACGCT 784
 Db 764 TCATCTTCTGTGCTCACCGCACTTCCCGTGGGAGCGCGTGGCGCGCGACGCT 823
 QY 785 TCTTCAGAGTTCGTGCGTGGCAGCGGGCGCGCTCCCGGGCTGCTTCGCAAGTGGC 844
 Db 824 TCTTCAGAGTTCGTGCGTGGCAGCGGGCGCGCTCCCGGGCTGCTTCGCAAGTGGC 883
 QY 845 GCCGCTTACAGAGCGCGCTGGCGCATGTTCCAGGCTTACTGCGCTTGGAGCCCGGAGC 904

Db 884 GCCGCTTACCGAGCCCGCTGCGCATGTTCCAGCGTTACTGCGCTTGGAGCCCGGAGC 943
 QY 905 GCGCGGCCCGAGCCAAAGGAGGTGTTCCCGTTTCTCAAGCACGAGTCACTCCGAGCTGC 964
 Db 944 GCGCGGCCCGAGCCAAAGGAGGTGTTCCCGTTTCTCAAGCACGAGTCACTCCGAGCTGC 1003
 QY 965 GCGCGGCCCGCTTCGACCGCGCGGCAAGCCCGCGGGGAGCCCGCGCGCGCGCGGC 1024
 Db 1004 GCGCGGCCCGCTTCGACCGCGCGGCAAGCCCGCGGGGAGCCCGCGCGCGCGCGGC 1063
 QY 1025 CACTGCGCTCGAGCGCGCTGAGCGGCTCAAGGGACGCTGCTGACCGAGAGCGGCGAGC 1084
 Db 1064 CACTGCGCTCGAGCGCGCTGAGCGGCTCAAGGGACGCTGCTGACCGAGAGCGGCGAGC 1123
 QY 1085 GTTCCCGCGCGCGCGCGCGCGCGCTGCGGTGCGCTTGCCTTGCCTGCGCGCTGC 1144
 Db 1124 GCTCCCGCGCGCGCGCGCGCGCGCTGCGGTGCGGTGCGCTTGCCTGCGCGCTGC 1183
 QY 1145 CAGTGCCTGCGCGGTGCTGCGCGAGCCCGCGCTAGCTCCCGAGGGGCGCGCGCGGC 1204
 Db 1184 CAGTGCCTGCGCGGTGCTGCGCGAGCCCGCGCTAGCTCCCGAGGGGCGCGCGCGGC 1243
 QY 1205 GGACCGACGCGCGCGGCAAGAGCAGAGGAGGAGGAGTGTGTGCGCACGCGCATCGAG 1264
 Db 1244 GGACCGACGCGCGCGGCAAGAGCAGAGGAGGAGGAGTGTGTGCGCACGCGCATCGAG 1303
 QY 1265 TCTGCGTCTGA 1275
 Db 1304 TCTGCGTCTGA 1314
 RESULT 9
 AAA09825
 ID AAA09825 standard; DNA; 1527 BP.
 XX
 AC AAA09825;
 XX 05-JUL-2000 (first entry)
 XX Rat PKS gene.
 XX PKS protein; Wistar rat; cerebral nervous system disease;
 XX neurological function-related disease; ds.
 XX Rattus sp.
 XX JP2000060571-A.
 XX 29-FEB-2000.
 XX 20-AUG-1998; 98JP-0249064.
 XX 20-AUG-1998; 98JP-0249064.
 XX (MITU) MITSUBISHI CHEM CORP.
 XX WPI; 2000-249682/22.
 XX P-PSDB; AAY83904.
 XX Novel mammalian peptide and a polynucleotide encoding it - useful for
 XX treatment and diagnosis of cerebral nervous system diseases and
 XX neurological function-related diseases
 XX Claim 15; Page 10-11; 15pp; Japanese.
 XX This sequence represents the gene encoding the PKS protein from Wistar
 XX rats. The peptide is useful for treatment and diagnosis of cerebral
 XX nervous system diseases and neurological function-related diseases.
 XX Sequence 1527 BP; 263 A; 495 C; 501 G; 268 T; 0 other;
 Query Match 75.6%; Score 963.4; DB 21; Length 1527;

181 AAAGGCACTATGGAAGGTTGACCTGGTGTCTACAAAGGCACAGGCACAAAAATGSCA 240
Db |||||
181 AAAGGCACTATGGAAGGTTGACCTGGTGTCTACAAAGGCACAGGCACAAAAATGSCA 240
QY
241 CTGAAGTTTGTGAACAGAGCAACCAAGCTGAGACTTCCCTACGGAGGTGAGCATC 300
Db |||||
241 CTGAAGTTTGTGAACAGAGCAACCAAGCTGAGAACTTCCCTACGGAGGTGAGCATC 300
QY
301 ACCAACAGCCCTCTCCTCCAGGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTTGAGACA 360
Db |||||
301 ACCAACAGCCCTCTCCTCCAGGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTTGAGACA 360
QY
361 GAGACTGCTACGTTCTTTCGCCAGGAGTACGACCTGCTGGGACCTTTTGACATATC 420
Db |||||
361 GAGACTGCTACGTTCTTTCGCCAGGAGTACGACCTGCTGGGACCTTTTGACATATC 420
QY
421 CCTCCAGGTGGGCTCCCTGAGACACGCTGAGCGCTGTGTGACAGCTGGGCTG 480
Db |||||
421 CCTCCAGGTGGGCTCCCTGAGACACGCTGAGCGCTGTGTGACAGCTGGGCTG 480
QY
481 CGCTGTGACTTCATGCAAGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAACCTG 540
Db |||||
481 CGCTGTGACTTCATGCAAGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAACCTG 540
QY
541 CTGCTTTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGACTTCGGCATGACGCGCGCG 600
Db |||||
541 CTGCTTTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGACTTCGGCATGACGCGCGCG 600
QY
601 GTGGCTCCCGCTCAAGCGGTGAGCGGACCATTCCTTACAGCGCGCTGAGTGTGC 660
Db |||||
601 GTGGCTCCCGCTCAAGCGGTGAGCGGACCATTCCTTACAGCGCGCTGAGTGTGC 660
QY
661 CAGCGCGCGCGCGCGCTGGCGGTGGACACAGCGCGCTGGACGTGTGGGCGCTTCGGC 720
Db |||||
661 CAGCGCGCGCGCGCGCTGGCGGTGGACACAGCGCGCTGGACGTGTGGGCGCTTCGGC 720
QY
721 GTGCTCATCTTCTGCGTGTCTACCGGCACTTCCCGTGGAGGCGCGCTGCGGGCGCGAC 780
Db |||||
721 GTGCTCATCTTCTGCGTGTCTACCGGCACTTCCCGTGGAGGCGCGCTGCGGGCGCGAC 780
QY
781 GCCTTTTCGAGGAGTTCGTGCGTGGCGAGCGGGCGCGCTGCGGGCGCTGCGTTCGAG 840
Db |||||
781 GCCTTTTCGAGGAGTTCGTGCGTGGCGAGCGGGCGCGCTGCGGGCGCTGCGTTCGAG 840
QY
841 TGGCGCGCTTCACCGAGCGCGCTGCGCATGTTTCCAGCGCTTACTGGCGCTTGAGCGCC 900
Db |||||
841 TGGCGCGCTTCACCGAGCGCGCTGCGCATGTTTCCAGCGCTTACTGGCGCTTGAGCGCC 900
QY
901 GAGCGCGCGCGCGCTGAGAGGTGTTCGCTTCTCAAGCAGAGCTACGTTCCGAG 960
Db |||||
901 GAGCGCGCGCGCGCTGAGAGGTGTTCGCTTCTCAAGCAGAGCTACGTTCCGAG 960
QY
961 CTGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db |||||
961 CTGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY
1021 GGGCACTCGGCTCGAGCGCGCTGGCGCTTCAAGCGGAGCGGTGCTACCGAGAGCGCG 1080
Db |||||
1021 GGGCACTCGGCTCGAGCGCGCTGGCGCTTCAAGCGGAGCGGTGCTACCGAGAGCGCG 1080
QY
1081 AGCGCTTCGCGCGCGCGCGCGCTGCGGTGCGTTCCTTCCGCTGCGGTGCGGTGCGG 1140
Db |||||
1081 AGCGCTTCGCGCGCGCGCGCGCTGCGGTGCGTTCCTTCCGCTGCGGTGCGGTGCGG 1140
QY
1141 GTGCCAGTCCGCTGCGTGTGCGCGAGCGCGCTAGCTCCCGAGGGCGCGCGCGCG 1200
Db |||||
1141 GTGCCAGTCCGCTGCGTGTGCGCGAGCGCGCTAGCTCCCGAGGGCGCGCGCGCGCG 1200
QY
1201 GSCCGGACCGAGCGCGCGCGGACAAAGGCGAGGTGTGTGCTGCCACGCGCCATC 1260
Db |||||
1201 GSCCGGACCGAGCGCGCGCGGACAAAGGCGAGGTGTGTGCTGCCACGCGCCATC 1260

Qy 1261 GAGATCTGCCTC 1272
Db |||||
1261 GAGATCTGCCTC 1272
RESULT 8
ABK10101
ID ABK10101 standard; cDNA; 2893 BP.
XX ABK10101;
AC
XX
DT 21-MAY-2002 (first entry)
XX
DE Human cDNA encoding protein kinase 32374.
XX
KW Human; ss; gene; protein kinase; 32374; 18431; cancer; infection;
cellular proliferation; pain; inflammation; metabolic disorder; obesity;
lipid disorder; diabetes; brain disorder; Alzheimer's disease;
neural tube defect; cerebrovascular disease; hypoxia; AIDS;
acquired immunodeficiency syndrome associated myopathy; heart disorder;
multiple sclerosis; angina pectoris; blood vessel disorder; vasculitis;
bone disorder; osteoporosis; haematopoietic disorder; autoimmune disease;
rheumatoid arthritis; osteoarthritis; multiple sclerosis;
systemic lupus erythematosus; asthma; Grave's disease; liver disorder;
graft-versus-host disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 274..1314
FT /**tag= a
FT /product= "Protein kinase 32374"
XX
PN WO200210401-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-US23653.
XX
PR 28-JUL-2000; 2000US-221543P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PI Meyers R, Kapeller-Libermann R, Silos-Santiago I;
XX
XX WPI; 2002-188748/24.
XX
XX P-PSDB; AAU76323.
XX
XX Novel protein kinase nucleic acid molecules and the encoded proteins
for diagnosing and treating cellular proliferative, bone, immune,
cardiovascular, liver, pain or metabolic disorders and identifying
modulators -
XX
XX Claim 1; Fig 1; 141pp; English.
XX
XX The invention relates to an isolated protein kinase polypeptide termed as
32374 or 18431, their allelic variants, fragments and nucleic acids
encoding them. Also included are a host cell containing the nucleic
acid, an antibody that binds to the protein kinase and modulators
of the activity of the kinase. The proteins, nucleic acids,
antibodies and modulators are useful for diagnosing, treating or
monitoring cancer or aberrant cellular proliferation and/or
differentiation such as lung, ovarian or brain cancer, pain (e.g.
inflammation or infection) or metabolic disorder (e.g. hypoxia,
ischemia, infarction), AIDS (acquired immunodeficiency syndrome) -
disease), neural tube defects, cerebrovascular diseases (e.g. Huntington's
disease), cachexia, lipid disorders and diabetes) or brain disorder
(e.g. Alzheimer's disease, Pick's disease, Parkinsonism, Huntington's
disease), associated myopathy, multiple sclerosis, disorders of the heart
(e.g. heart failure, angina pectoris, myocardial infarction),
blood vessel disorders (e.g. arteriovenous fistula, vasculitis,
syphilitic aneurysms), disorders of the bone (e.g. osteoporosis,
osteitis, hypoparathyroidism, hyperparathyroidism, tropical sprue,


```
QY 421 CTTCCAGGTGGGCTCCGTGAGACACCGTGAAGCGTGTGTGACGAGCTGGGCGTG 480
Db 777 CTTCCAGGTGGGCTCCGTGAGACACCGTGAAGCGTGTGTGACGAGCTGGGCGTG 836
QY 481 GCGTGCATTCATGACGCGGCGGAGCTGCTGTCACCGGACATCAAGCCCGAGACGTG 540
Db 837 GCGTGCATTCATGACGCGGCGGAGCTGCTGTCACCGGACATCAAGCCCGAGACGTG 896
QY 541 CTGCTGTTACGCGGAGTCCGCGGCTAAAGTGGCCGACTTCGSCATGACGCGCGC 600
Db 897 CTGCTGTTACGCGGAGTCCGCGGCTAAAGTGGCCGACTTCGSCATGACGCGCGC 956
QY 601 GTGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGGCGCTGAGGTGTC 660
Db 957 GTGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGGCGCTGAGGTGTC 1016
QY 661 CAGCGGCGCGCGCGGAGTGGCGGTGACACGCGGCTGACGCTGTGGGCTTCGCG 720
Db 1017 CAGCGGCGCGCGCGGAGTGGCGGTGACACGCGGCTGACGCTGTGGGCTTCGCG 1076
QY 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGTGGAGGCGGCTCGGCGCGGAC 780
Db 1077 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGTGGAGGCGGCTCGGCGCGGAC 1136
QY 781 GCCTTCTTGAAGAGTTCTGCGCTGACGCGGCGGCGGCTGCGGCGGCTGCTTCGAG 840
Db 1137 GCCTTCTTGAAGAGTTCTGCGCTGACGCGGCGGCGGCTGCGGCGGCTGCTTCGAG 1196
QY 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCGCTTGGAGCCC 900
Db 1197 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCGCTTGGAGCCC 1256
QY 901 GAGCGCGCGCGCGCGAGGAGTTCGCTTCTCAAGCACGAGCTCAAGTTCGAG 960
Db 1257 GAGCGCGCGCGCGCGAGGAGTTCGCTTCTCAAGCACGAGCTCAAGTTCGAG 1316
QY 961 CTGGCGCGCGCGCTGCGACCGCGCGGCGGACGCGCGCGCGCGCGCGCGCGCG 1020
Db 1317 CTGGCGCGCGCGCTGCGACCGCGCGGCGGACGCGCGCGCGCGCGCGCGCGCG 1376
QY 1021 GGGCCACTGCGCTCGAGGCGCTGGCGCGCTCAAGCGGAGCGGTGCTGACGAGCGGC 1080
Db 1377 GGGCCACTGCGCTCGAGGCGCTGGCGCGCTCAAGCGGAGCGGTGCTGACGAGCGGC 1436
QY 1081 AGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db 1437 GCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1496
QY 1141 GTGCCAGTCCCGTCCCGTCCCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCG 1200
Db 1497 GTGCCAGTCCCGTCCCGTCCCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCG 1556
QY 1201 GCGCGGACCGAGCGCGCGCGGACAAAGGAGGAGTGTGTGCGCGCGCGCGCGCG 1260
Db 1557 GCGCGGACCGAGCGCGCGCGGACAAAGGAGGAGTGTGTGCGCGCGCGCGCGCG 1616
QY 1261 GAGATCTGCGTCTGA 1275
Db 1617 GAGATCTGCGTCTGA 1631
```

RESULT 7

```
AAD49416
ID AAD49416 standard; DNA; 1272 BP.
XX
AC AAD49416;
XX
DT 24-MAR-2003 (first entry)
XX Human serine/threonine protein kinase-like protein (STPKP) DNA #1.
DE Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
XX
```

chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1272
FT /tag= a
FT /product= "Human STPKP protein"
FT /note= "CDS does not include stop codon"
FT /partial

W0200283882-A2.

24-OCT-2002.

12-APR-2002; 2002WO-BP04080.

12-APR-2001; 2001US-283189P.

04-JUN-2001; 2001US-294989P.

30-JUL-2001; 2001US-308094P.

08-FEB-2002; 2002US-354574P.

(FARB) BAYER AG.

Koehler RH;

WPI; 2003-093018/08.

P-PSDB; AAE32010.

New polynucleotides encoding serine/threonine protein kinase-like
protein (STPKP) and its encoded protein, useful for identifying
modulators of STPKP activity, and in gene therapy for treating e.g.
cancer, obesity or diabetes.

Claim 1; Page 146-147; 161pp; English.

The invention relates to human serine/threonine protein kinase-like
protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
identifying test compounds, that may act as agonists or antagonists at
the receptor site and which can be regulated to provide therapeutic
effects. The reagent (e.g. antibody or antisense oligonucleotide, which
decreases the expression of human STPKP gene or decreases the levels of
STPKP protein), STPKP agonist, STPKP protein or expression vector is
useful for treating a patient with a chronic obstructive pulmonary
disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
disease, diabetes, obesity or urology disorders. The vectors comprising
STPKP DNA and reagents are useful for preparing a medicament for
modulating the activity of STPKP in the diseases. These diseases include
multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
urinary incontinence, carcinoma, leukaemia, or benign prostatic
hypoplasia. STPKP gene is used in gene therapy. The present sequence is
human STPKP DNA.

Sequence 1272 BP; 191 A; 450 C; 430 G; 201 T; 0 other;

Query Match 99.0%; Score 1262.4; DB 25; Length 1272;

Best Local Similarity 99.5%; Pred. No. 1.3e-202;

Matches 1266; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 60

Db 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 60

QY 61 ACTGCCCTGGGCTGTGTCGGTGTGCCCTTCTCACTGAACATGAGGCGCTGACT 120

Db 61 ACTGCCCTGGGCTGTGTCGGTGTGCCCTTCTCACTGAACATGAGGCGCTGACT 120

QY 121 CTCGGCACACTGGCGCGCGAGCGACGTGACCAAGCACTACGAAGTGTCCGGGAGCTGGGC 180

Db 121 CTCGGCACACTGGCGCGCGAGCGACGTGACCAAGCACTACGAAGTGTCCGGGAGCTGGGC 180


```
Db 385 CTGCTGTTGACCCGAGTCCGCGCGTAAAGCTGGCCGACTTCGGATGACCGCGCC 1044
Qy 601 GTGGGTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGCGCTTGAAGTGTGC 660
Db 1045 GTGGGTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGCGCTTGAAGTGTGC 1104
Qy 661 CAGCGGCGCGCGCGCGCGCTGCGGTGACACGGCGGTGACGCTGTGGCGCTTGGC 720
Db 1105 CAGCGGCGCGCGCGCGCGCTGCGGTGACACGGCGGTGACGCTGTGGCGCTTGGC 1164
Qy 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGCTGGGAGGCGCGCTCGGGCGCCGAC 780
Db 1165 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGCTGGGAGGCGCGCTCGGGCGCCGAC 1224
Qy 781 GCCTTCTCGAGAGTGTGCTGCTGCGTGAAGCGGCGCGCTTCCGCGGCTGCTTGGCGAG 840
Db 1225 GCCTTCTCGAGAGTGTGCTGCTGCGTGAAGCGGCGCGCTTCCGCGGCTGCTTGGCGAG 1284
Qy 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCGCTGAGGCC 900
Db 1285 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCGCTGAGGCC 1344
Qy 901 GAGCGCGCGCGCGCGCGAGTGTTCGCTTCTTCAAGCAGAGCTCACGCTCCGAG 960
Db 1345 GAGCGCGCGCGCGCGCGAGTGTTCGCTTCTTCAAGCAGAGCTCACGCTCCGAG 1404
Qy 961 CTGGCGCGCGCGCTTGCACCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 1405 CTGGCGCGCGCGCTTGCACCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1464
Qy 1021 GGGCACTCGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGAGCGTGTGACCGAGCGCG 1080
Db 1465 GGGCACTCGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGAGCGTGTGACCGAGCGCG 1524
Qy 1081 AGCGGCTCCCGCGCGCGCGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG 1140
Db 1525 GCGGCTCCCGCGCGCGCGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG 1584
Qy 1141 GTGCGAGTCCCGTGGCGCGCTGTGCGCGAGCGCGCGCTAGCTCCCGAGCGCGCG 1200
Db 1585 GTGCGAGTCCCGTGGCGCGCTGTGCGCGAGCGCGCGCTAGCTCCCGAGCGCGCG 1644
Qy 1201 GCGCGGACCGCGCGCGCGCGCAAGGCAAGGCGAGGTGTGCTGGCGCACCGCGCATC 1260
Db 1645 GCGCGGACCGCGCGCGCGCGCAAGGCAAGGCGAGGTGTGCTGGCGCACCGCGCATC 1704
Qy 1261 GAGATCTGCGTCTGA 1275
Db 1705 GAGATCTGCGTCTGA 1719

RESULT 6
AAD49420
ID AAD49420 standard; DNA; 3124 BP.
AC AAD49420;
XX
XX
XX
XX 24-MAR-2003 (first entry)
DE Human serine/threonine protein kinase-like protein (STPKP) DNA #5.
KW Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
KW anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.
OS Homo sapiens.
XX
XX
XX WO200283882-A2.
XX
```

```
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-EP04080.
XX
PR 12-APR-2001; 2001US-283189P.
PR 04-JUN-2001; 2001US-294989P.
PR 30-JUL-2001; 2001US-308094P.
PR 08-FEB-2002; 2002US-354574P.
XX
PA (FARB ) BAYER AG.
XX
XX Koehler RH;
XX
XX WPI; 2003-093018/08.
XX
XX New polynucleotides encoding serine/threonine protein kinase-like
PT protein (STPKP) and its encoded protein, useful for identifying
PT modulators of STPKP activity, and in gene therapy for treating e.g.
PT cancer, obesity or diabetes .
XX
XX Claim 1; Page 159-161; 161pp; English.
XX
XX The invention relates to human serine/threonine protein kinase-like
CC protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
CC identifying test compounds, that may act as agonists or antagonists at
CC the receptor site and which can be regulated to provide therapeutic
CC effects. The reagent (e.g. antibody or antisense oligonucleotide, which
CC decreases the expression of human STPKP gene or decreases the levels of
CC STPKP protein), STPKP agonist, STPKP protein or expression vector is
CC useful for treating a patient with a chronic obstructive pulmonary
CC disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
CC disease, diabetes, obesity or urology disorders. The vectors comprising
CC STPKP DNA and reagents are useful for preparing a medicament for
CC modulating the activity of STPKP in the diseases. These diseases include
CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
CC human STPKP DNA.
XX
XX Sequence 3124 BP; 485 A; 1110 C; 1081 G; 448 T; 0 other;
SQ
Query Match 99.9%; Score 1273.4; DB 25; Length 3124;
Best Local Similarity 99.9%; Pred. No. 2e-204;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGCGTGGCTGCGCCAGAGCGCTGAGCGCGCGCGCTGAGCGCGCTGCTGCGCGCGCGG 60
Db 357 ATGAGCGTGGCTGCGCCAGAGCGCTGAGCGCGCGCGCTGAGCGCGCTGCTGCGCGCGCGG 416
Qy 61 ACTGCCCTGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGCT 120
Db 417 ACTGCCCTGGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGCT 476
Qy 121 CTCGCGACACTGGCGCGCGCGCGCGCTGAGCGCGCGCGCGCTGAGCGCGCGCGCGCGCGCG 180
Db 477 CTCGCGACACTGGCGCGCGCGCGCGCGCTGAGCGCGCGCGCGCTGAGCGCGCGCGCGCGCG 536
Qy 181 AAGGCGACCTATGGGAAGGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 537 AAGGCGACCTATGGGAAGGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
Qy 241 CTGAAGTTTGTGAACAGAGCAAAACCAAGCTGGAAGCACTTCTTACGGAGGTGAGCATC 300
Db 597 CTGAAGTTTGTGAACAGAGCAAAACCAAGCTGGAAGCACTTCTTACGGAGGTGAGCATC 656
Qy 301 ACCAACAGCTCTCTCTCCAGCGCGCGCTTCTCATCATCAAGTCTTTGACGTGCTTTTGAACA 360
Db 657 ACCAACAGCTCTCTCTCCAGCGCGCGCTTCTCATCATCAAGTCTTTGACGTGCTTTTGAACA 716
Qy 361 GAGGACTGTGCTGCTTTTGGCGCGAGTACGACCTGCTGGGAGCTGTTTGTGACATCATC 420
Db 717 GAGGACTGTGCTGCTTTTGGCGCGAGTACGACCTGCTGGGAGCTGTTTGTGACATCATC 776
```


Db 1021 GGGCCACTGCGCTCGAGCGCCCTGGCGCGCTCAAGCGGACGGTGTGACCGGAGAGCGGC 1080
Qy 1081 AGCGCTCCCGCGCCCGCGCCCGCGCTCGGTCGGTGCCTTGCCTTGCCTGCGGTCGCG 1140
Db 1081 GCGGCTCCCGCGCCCGCGCCCGCGCTCGGTCGGTGCCTTGCCTTGCCTGCGGTCGCG 1140
Qy 1141 GTGCCAGTGCCTGCGGTCCTGTGCGCGAGCCGCGCTAGTCTCCCGAGGCGCCGCC 1200
Db 1141 GTGCCAGTGCCTGCGGTCCTGTGCGCGAGCCGCGCTAGTCTCCCGAGGCGCCGCC 1200
Qy 1201 GCGCGACCGACGCGCGCGGACAGAGCAAGAGCAAGAGGAGGTCGTCGCGCCACGSCCATC 1260
Db 1201 GCGCGACCGACGCGCGCGGACAGAGCAAGAGGAGGTCGTCGCGCGAGGTCGTCGCGCCACGSCCATC 1260
Qy 1261 GAGATCTGCGTCTGA 1275
Db 1261 GAGATCTGCGTCTGA 1275

RESULT 4

AAD49419
ID AAD49419 standard; DNA; 1275 BP.
AC AAD49419;
DT 24-MAR-2003 (first entry)
DE Human serine/threonine protein kinase-like protein (STPKP) DNA #4.
KW Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
KW anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT 1..1275
XX CDS /*tag= a
XX FT /*product= "Human STPKP protein"
XX
XX WO200283882-A2.
XX
XX PD 24-OCT-2002.
XX
XX PF 12-APR-2002; 2002WO-EP04080.
XX
XX PR 12-APR-2001; 2001US-283189P.
XX PR 04-JUN-2001; 2001US-294989P.
XX PR 30-JUL-2001; 2001US-308094P.
XX PR 08-FEB-2002; 2002US-354574P.
XX
XX PA (FARB) BAYER AG.
XX
XX PI Koehler RH;
XX
XX DR WPI; 2003-093018/08.
XX DR P-PSDB; AAE32011.
XX
XX PT New polynucleotides encoding serine/threonine protein kinase-like
XX PT protein (STPKP) and its encoded protein, useful for identifying
XX PT modulators of STPKP activity, and in gene therapy for treating e.g.
XX PT cancer, obesity or diabetes -
XX
XX PS Claim 1; Page 156-157; 161pp; English.

CC The invention relates to human serine/threonine protein kinase-like
CC protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
CC identifying test compounds, that may act as agonists or antagonists at
CC the receptor site and which can be regulated to provide therapeutic
CC effects. The reagent (e.g. antibody or antisense oligonucleotide, which

CC decreases the expression of human STPKP gene or decreases the levels of
CC STPKP protein), STPKP agonist, STPKP protein or expression vector is
CC useful for treating a patient with a chronic obstructive pulmonary
CC disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
CC disease, diabetes, obesity or urology disorders. The vectors comprising
CC STPKP DNA and reagents are useful for preparing a medicament for
CC modulating the activity of STPKP in the diseases. These diseases include
CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
CC human STPKP DNA.
XX

SQ Sequence 1275 BP; 191 A; 450 C; 434 G; 200 T; 0 other;

Query Match 99.9%; Score 1273.4; DB 25; Length 1275;

Best Local Similarity 99.9%; Pred. No. 1.9e-204;

Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGCGTGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTGTGGCCCGGG 60
Db 1 ATGAGCGTGGCTGCCAGAGCCTGAGCCGCCCGCTCCCTGACCTCTGTGGCCCGGG 60
Qy 61 ACTGCCCTGGCGCTGTCGCGGTGTCCTCACTGAAGACATGACAGCCCTGACT 120
Db 61 ACTGCCCTGGCGCTGTCGCGGTGTCCTCACTGAAGACATGACAGCCCTGACT 120
Qy 121 CTCGCGACACTGGCCGCCAGCGAGCTCACCAGCACTACGAACCTAGTCCGGAGCTGGC 180
Db 121 CTCGCGACACTGGCCGCCAGCGAGCTCACCAGCACTACGAACCTAGTCCGGAGCTGGC 180
Qy 181 AAGGCACTTATGGAGAGTTGACCTGTGTGTACAGGCGACAGCAAAAATGGCA 240
Db 181 AAGGCACTTATGGAGAGTTGACCTGTGTGTACAGGCGACAGCAAAAATGGCA 240
Qy 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
Db 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
Qy 301 ACCAAGCCTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTTTGAGACA 360
Db 301 ACCAAGCCTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTTTGAGACA 360
Qy 361 GAGGACTGTCTACGTCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTGACATATC 420
Db 361 GAGGACTGTCTACGTCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTGACATATC 420
Qy 421 CCTCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGTGTGTGACGAGCTGGGCTG 480
Db 421 CCTCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGTGTGTGACGAGCTGGGCTG 480
Qy 481 GCGTGTGACTTATGCAAGCGGCGGAGTGTGTGACCGGACATCAAGCCGAGAACTG 540
Db 481 GCGTGTGACTTATGCAAGCGGCGGAGTGTGTGACCGGACATCAAGCCGAGAACTG 540
Qy 541 CTGCTGTTCGACCGGAGTGGCGCGTAAAGTGGCGGACCTTCGCGCATGACCGCGGC 600
Db 541 CTGCTGTTCGACCGGAGTGGCGCGTAAAGTGGCGGACCTTCGCGCATGACCGCGGC 600
Qy 601 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGCGCTGAGGTGTC 660
Db 601 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGCGCTGAGGTGTC 660
Qy 661 CAGCGGCGCGCGCGGCTGGCGTGGACACCGGCGCTGGACACCGGCGCTGGGCTTCGGC 720
Db 661 CAGCGGCGCGCGCGGCTGGCGTGGACACCGGCGCTGGACACCGGCGCTGGGCTTCGGC 720
Qy 721 GTGCTCATCTTCGCTGCTCACCGGCAACTTCCGTTGGAGGCGGCTGGGCGCGGCAC 780
Db 721 GTGCTCATCTTCGCTGCTCACCGGCAACTTCCGTTGGAGGCGGCTGGGCGCGGCAC 780
Qy 781 GCCTTCTTCAGAGGATTTGTCGCTGCTGACGCGGCGGCTGCGCGGCGCTGCCCTTCGAG 840
Db 781 GCCTTCTTCAGAGGATTTGTCGCTGCTGACGCGGCGGCTGCGCGGCGCTGCCCTTCGAG 840

QY 1141 GTCCCAATGCTCCGCTGCGGTGCTGTGCGAGCCCGGCTAGCTCCCGAGGGCCCCC 1200
 Db 1914 GTCCCAATGCTCCGCTGCGGTGCTGTGCGAGCCCGGCTAGCTCCCGAGGGCCCCC 1973
 QY 1201 GCGCGACGACGCGCGCGGACACAGCAAGAGGACAGTGTGTGGCCACGGGCATC 1260
 Db 1974 GCGCGACGACGCGCGCGGACACAGCAAGAGGACAGTGTGTGGCCACGGGCATC 2033
 QY 1261 GAGATCTGCTCTGA 1275
 Db 2034 GAGATCTGCTCTGA 2048
 RESULT 3
 AAS06744
 ID AAS06744 standard; cDNA; 1275 BP.
 AC AAS06744;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Polynucleotide sequence encoding human protein kinase #44.
 XX
 KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 DN W0200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32085.
 XX
 PR 24-NOV-1999; 99US-0167482.
 XX
 PA (SUGE-) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 DR P-PSDB; AAU03544.
 XX
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Example 1; Figure 1; 433pp; English.
 XX
 CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 SQ Sequence 1275 BP; 191 A; 450 C; 434 G; 200 T; 0 other;
 Query Match 99.9%; Score 1273.4; DB 22; Length 1275;

Best Local Similarity 99.9%; Pred. No. 1.9e-204;
 Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAGCGTGGGTGCTCCAGAGCTGAGCGCCCGCTCCCTGACCTGTGTGGCGGG 60
 Db 1 ATGAGCGTGGGTGCTCCAGAGCTGAGCGCCCGCTCCCTGACCTGTGTGGCGGG 60
 QY 61 ACTGCCCTTGGGCTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 61 ACTGCCCTTGGGCTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 121 CTCCGCACACTGCGCGCCAGCGACGTCACAAAGCACTACGAACTAGTCCGGAGCTGGGC 180
 Db 121 CTCCGCACACTGCGCGCCAGCGACGTCACAAAGCACTAGTCCGGAGCTGGGC 180
 QY 181 AAAGGCACCTATGGGAAGTGTGACTGTGGTGTCTCAAGGGCACAGGCAAAATGGCA 240
 Db 181 AAAGGCACCTATGGGAAGTGTGACTGTGGTGTCTCAAGGGCACAGGCAAAATGGCA 240
 QY 241 CTGAGTTTGTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGAGGTGACATC 300
 Db 241 CTGAGTTTGTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGAGGTGACATC 300
 QY 301 ACCAAGAGCTCTCCCTCCAGCCCTTCAATCATCAAGGTCCTTGAACCTGTCTTTGAGACA 360
 Db 301 ACCAAGAGCTCTCCCTCCAGCCCTTCAATCATCAAGGTCCTTGAACCTGTCTTTGAGACA 360
 QY 361 GAGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 GAGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 CCTCCCGAGTGGGCTCCCTGAGGACAGGTGAAGCGTGTGTGACAGCTGGGCTG 480
 Db 421 CCTCCCGAGTGGGCTCCCTGAGGACAGGTGAAGCGTGTGTGACAGCTGGGCTG 480
 QY 481 GCGCTGGAATTCATGCAACGGGCGGAGCTGTGTGCAACGGGACATCAAGCCGAGAACGTG 540
 Db 481 GCGCTGGAATTCATGCAACGGGCGGAGCTGTGTGCAACGGGACATCAAGCCGAGAACGTG 540
 QY 541 CTGCTGTTGACCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 541 CTGCTGTTGACCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 601 GTGGGCTGCGCGCTCAAGCGCTGAGCGGACCATCCCTTACACGGCGCTGAGGTGTGC 660
 Db 601 GTGGGCTGCGCGCTCAAGCGCTGAGCGGACCATCCCTTACACGGCGCTGAGGTGTGC 660
 QY 661 CAGCGCGCGCGCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 CAGCGCGCGCGCGCGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 GTGCTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 GTGCTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 GCTTCTTCTGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 GCTTCTTCTGAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 TGGCGCGCTTCAACCGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 841 TGGCGCGCTTCAACCGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 GAGCGCGCGCGCGCGAGAGAGTGTGCTGCTTCTCAAGCACAGGCTCAGTCCGAG 960
 Db 901 GAGCGCGCGCGCGCGAGAGAGTGTGCTGCTTCTCAAGCACAGGCTCAGTCCGAG 960
 QY 961 CTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 961 CTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 GGGGCACTGCGCTGAGGGGCTTGGGCGCTCAAGCGGAGCGGTGTGACCGAGAGCGGC 1080

PH	Key	Location/Qualifiers	
FT	5'UTR	1-774	
FT		/*tag= a	
FT	CDS	775..2048	
FT		/*tag= b	
FT	3'UTR	/product= "Protein Kinase N"	
FT		2049..2598	
FT		/*tag= c	
XX			
PN	WO200188148-A2.		
XX			
PD	22-NOV-2001.		
XX			
PF	17-MAY-2001; 2001WO-US15776.		
XX			
PR	17-MAY-2000; 2000US-205228P.		
PR	12-DEC-2000; 2000US-0734032.		
PR	26-MAR-2001; 2001US-0816094.		
XX			
PA	(APPL-) APPLERA CORP.		
XX			
PI	Wei M, Chandramouliswara I, Ye J, Ketchum KA, Di Francesco V;		
PI	Beasley EM;		
XX			
DR	P-PSDB; RAU10023.		
XX			
PT	Human kinase protein and polynucleotides encoding them, useful for		
PT	identifying modulators of kinase polypeptides and for treating,		
PT	preventing, and/or diagnosing neurodegenerative diseases and cancer -		
XX			
PS	Claim 4; Figure 1; 65pp; English.		
XX			
CC	This sequence represents a human protein kinase N cDNA of the invention.		
CC	The invention comprises nucleotide and protein sequences of an isolated		
CC	protein which is related to the PKN kinase subfamily. The protein		
CC	kinase N gene is located on chromosome 16. The protein may have		
CC	cytostatic and neuroprotective and can be used in gene therapy		
CC	possibly as a human kinase protein expression or activity modulator.		
CC	The nucleic acids and polypeptides of the invention may be used in the		
CC	prevention, diagnosis and treatment of diseases associated with		
CC	inappropriate kinase expression. The nucleic acids (or vectors		
CC	containing them) and the kinase may be used to treat disorders		
CC	associated with decreased expression, either by rectifying mutations or		
CC	deletions in a patient's genome that affect the activity of the enzyme		
CC	by expressing inactive proteins or to supplement the patients own		
CC	production of kinases. Additionally, the nucleic acids may be used to		
CC	produce the kinase, by inserting the nucleic acids into a host cell and		
CC	culturing the cell to express the protein. The nucleic acid and its		
CC	complementary sequences may also be used as DNA probes in diagnostic		
CC	assays to detect and quantitate the presence of similar nucleic acids		
CC	in samples, and therefore which patients may be in need of restorative		
CC	therapy. The polypeptides may also be used as antigens in the		
CC	production of antibodies against the kinase and in assays to identify		
CC	modulators of kinase expression and activity. The anti-kinase		
CC	antibodies and antagonists may also be used to down regulate expression		
CC	and activity. The anti-kinase antibodies may also be used as diagnostic		
CC	agents for detecting the presence of kinase polypeptides in samples		
CC	(e.g. by enzyme linked immunosorbent assay (ELISA)). Disorders that may		
CC	be prevented, diagnosed and/or treated by the above methods include, for		
CC	example neurodegenerative diseases, inflammation, arteriosclerosis		
CC	psoriasis, cancer, papilloma virus infection, Alzheimer's disease		
CC	and growth disorders.		
XX			
SQ	Sequence 2598 BP; 446 A; 952 C; 862 G; 338 T; 0 other;		
	Query Match 100.0%; Score 1275; DB 24; Length 2598;		
	Best Local Similarity 100.0%; Pred. No. 1.1e-204;		
	Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGTGTGGCGCGGG 60		
Db			
	774 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 833		

Qy	61	ACTGCCCTGGGCTGTGTGGTGTGTCGCCCTTCTCACTGAAGACATGCGAGGCCCTGACT	120
Db	834	ACTGCCCTGGGCTGTGTGGTGTGTCGCCCTTCTCACTGAAGACATGCGAGGCCCTGACT	893
Qy	121	CTCCGACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGC	180
Db	894	CTCCGACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGC	953
Qy	181	AAAGGCACCTATGGGAAGTTGACCTTGCTCTACAAGGGCACAGGCACAAAAATGGCA	240
Db	954	AAAGGCACCTATGGGAAGTTGACCTTGCTCTACAAGGGCACAGGCACAAAAATGGCA	1013
Qy	241	CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAGAACTTCTACGGGAGGTGAGCATC	300
Db	1014	CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAGAACTTCTACGGGAGGTGAGCATC	1073
Qy	301	ACCAACAGCCTCTCTCCAGCCCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	360
Db	1074	ACCAACAGCCTCTCTCCAGCCCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	1133
Qy	361	GAGGACTGTACGTCTTTTGGCCAGGAGTACGCACCTCTCGGGACCTGTGTTGACATCATC	420
Db	1134	GAGGACTGTACGTCTTTTGGCCAGGAGTACGCACCTCTCGGGACCTGTGTTGACATCATC	1193
Qy	421	CCTCCCCAGGTGGGGCTCCCTGAGGACACGGTGAAGGCTGTGTGACAGACTGGGCGTG	480
Db	1194	CCTCCCCAGGTGGGGCTCCCTGAGGACACGGTGAAGGCTGTGTGACAGACTGGGCGTG	1253
Qy	481	CGCGTGACATTATGACACGGCGGCGACCTGCTGACCGGACATCAAGCCGAGAACGTG	540
Db	1254	CGCGTGACATTATGACACGGCGGCGACCTGCTGACCGGACATCAAGCCGAGAACGTG	1313
Qy	541	CTGCTGTTCACCGCAGTCCCGCGGTAAAGCTGGCCGACTTTCGCATGACCGCGCGC	600
Db	1314	CTGCTGTTCACCGCAGTCCCGCGGTAAAGCTGGCCGACTTTCGCATGACCGCGCGC	1373
Qy	601	GTGGGCTGCCGCTCAAGCCGCTGAGCGGCACCATCCCTTACACGGCGCTGAGTGTGC	660
Db	1374	GTGGGCTGCCGCTCAAGCCGCTGAGCGGCACCATCCCTTACACGGCGCTGAGTGTGC	1433
Qy	661	CAGCGCGCGCGCGCGCTGGCGGTGACACAGCGCGCTGACGCTGTGGGCTTCGCG	720
Db	1434	CAGCGCGCGCGCGCGCTGGCGGTGACACAGCGCGCTGACGCTGTGGGCTTCGCG	1493
Qy	721	GTGCTCATCTTCTGCGTGTCAACCGCACTTCCCGTGGAGGCGCGCTCGGGCGCGAC	780
Db	1494	GTGCTCATCTTCTGCGTGTCAACCGCACTTCCCGTGGAGGCGCGCTCGGGCGCGAC	1553
Qy	781	GCCTTCTTCAGAGATTCTGCGCTGCGAGGGGCGCGCTCGGGGCTCCCTTGGCAG	840
Db	1554	GCCTTCTTCAGAGATTCTGCGCTGCGAGGGGCGCGCTCGGGGCTCCCTTGGCAG	1613
Qy	841	TGGCGCGCTTCAACCGACCGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTGGAGCCC	900
Db	1614	TGGCGCGCTTCAACCGACCGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTGGAGCCC	1673
Qy	901	GAGCGCGCGCGCGCGCAAGAGAGTGTTCGCTTCTTCAAGCACAGCTCAGCTTCGAG	960
Db	1674	GAGCGCGCGCGCGCGCAAGAGAGTGTTCGCTTCTTCAAGCACAGCTCAGCTTCGAG	1733
Qy	961	CTGGCGCGCGCGCGCTGACCGCGCGCGCAAGCCCCCGGGGACCGCCCGCGCGCGCC	1020
Db	1734	CTGGCGCGCGCGCGCTGACCGCGCGCGCAAGCCCCCGGGGACCGCCCGCGCGCGCC	1793
Qy	1021	GGGCACTGCGCTCGAGGCGCTGGGCGCTCAAGCGGACGCTGTGTGACCGAGAGCGGC	1080
Db	1794	GGGCACTGCGCTCGAGGCGCTGGGCGCTCAAGCGGACGCTGTGTGACCGAGAGCGGC	1853
Qy	1081	AGCGGCTCCCGCGCGCGCGCGCGCTCGGGTGCCTTGGCCCTTGGCCGCTGCGCGTCCG	1140
Db	1854	AGCGGCTCCCGCGCGCGCGCGCGCTCGGGTGCCTTGGCCCTTGGCCGCTGCGCGTCCG	1913

XX Walke DW, Maricar M, Yu X, Friddle CJ;
XX WPI; 2002-527921/56.
XX P-PSDB; AAU79585.
XX Novel nucleic acid molecule encoding a human kinase, useful in
PT therapeutic, diagnostic and pharmacogenomic applications, as DNA
PT markers for restriction fragment length polymorphism analysis and in
PT forensic biology -
XX
PS Disclosure; Page 37; 37pp; English.
XX
CC The invention discloses an isolated human kinase polypeptide, designated
CC novel human protein (NHP), and the polynucleotide encoding it. Kinases
CC mediate phosphorylation of a wide variety of proteins and compounds in
CC the cell and are also involved in a range of regulatory pathways. The
CC polynucleotide is useful in therapeutic, diagnostic and pharmacogenomic
CC applications, and for identifying compounds that modulate, i.e. act as
CC agonists or antagonists of, the gene expression or gene product
CC activity. The polynucleotide is also useful as a probe in microarrays
CC or other cloning and/or assay formats, for screening collections of
CC genetic material from patients who have a particular medical condition,
CC for identifying mutations associated with a particular disease and also
CC as a diagnostic or prognostic assay. It can also be useful for the
CC detection of mutant human proteins, or inappropriately expressed
CC proteins, for the diagnosis of disease, for screening for drugs effective
CC in perturbing the normal function of the protein, for generation of
CC antibodies and as reagents in assays for screening for compounds that can
CC be used as pharmaceutical agents in the therapeutic treatment of mental,
CC biological or medical disorders and diseases. The sequence presented is
CC the novel human protein (NHP) cDNA which was isolated from human brain
CC and skeletal muscle cDNA libraries.
XX
SQ Sequence 1473 BP; 215 A; 533 C; 509 G; 216 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 1275; DB 24; Length 1473;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGTGGGTGCGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
DB 117 ATGAGCGTGGGTGCGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 176
QY 61 ACTGCCCTTGGCGCTGTGGTGTGGCGCTTCTCACTGAAGACATGAGCGCTGACT 120
DB 177 ACTGCCCTTGGCGCTGTGGTGTGGCGCTTCTCACTGAAGACATGAGCGCTGACT 236
QY 121 CTCGCGACACTGGCGCGCCAGCGACCTCACCAGCACTACGAACCTAGTCCGGAGCTGGGC 180
DB 237 CTCGCGACACTGGCGCGCCAGCGACCTCACCAGCACTAGTCCGGAGCTGGGC 296
QY 181 AAAGGCACTATGGGAAGTTGACCTGTGGTGTCTCAAGGGCACAGGCACAAAATGGCA 240
DB 297 AAAGGCACTATGGGAAGTTGACCTGTGGTGTCTCAAGGGCACAGGCACAAAATGGCA 356
QY 241 CTGAGTGTGTGAACAAGCAAAACCAAGCTGAGAACTTCTTACGGGAGTGAGCATC 300
DB 357 CTGAGTGTGTGAACAAGCAAAACCAAGCTGAGAACTTCTTACGGGAGTGAGCATC 416
QY 301 ACCAAGAGCTCTCTCCAGCGCCCTTCATCATCAAGGTCTTTGAGTGTCTTTGAGACA 360
DB 417 ACCAAGAGCTCTCTCCAGCGCCCTTCATCATCAAGGTCTTTGAGTGTCTTTGAGACA 476
QY 361 GAGGACTGCTACCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
DB 477 GAGGACTGCTACCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 536
QY 421 CCTCCCGAGTGGGCTCCCTGAGACACAGGTGAAGCGTGTGTGACAGCTGGGCGCTG 480
DB 537 CCTCCCGAGTGGGCTCCCTGAGACACAGGTGAAGCGTGTGTGACAGCTGGGCGCTG 596
QY 481 GCGCTGGACTTCATCAGCGGGCGGAGCTGTGGTGCACCGGACATCAAGCCCGAGACGTG 540

DB 597 GCGCTGACATTCATGCACGGGCGGACGTGGTGCACCGGACATCAAGCCCGAGAACGTG 656
QY 541 CTGCTGTTTCGACCGCGAGTGCCTGCGGTAAGCTGAGCGGCTTCGGCATCAGCGCGCG 600
DB 657 CTGCTGTTTCGACCGCGAGTGCCTGCGGTAAGCTGAGCGGCTTCGGCATCAGCGCGCG 716
QY 601 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGGCGCTGAGGTGTC 660
DB 717 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGGCGCTGAGGTGTC 776
QY 661 CAGCGGGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 720
DB 777 CAGCGGGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 836
QY 721 GTGCTCATCTTCTGCTGCTCACCAGGCTACTTCCGCTGGGAGGCGGCTGCGGCGCGGAC 780
DB 837 GTGCTCATCTTCTGCTGCTCACCAGGCTACTTCCGCTGGGAGGCGGCTGCGGCGCGGAC 896
QY 781 GCCTTCTTCGAGGAGTTCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 840
DB 897 GCCTTCTTCGAGGAGTTCGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 956
QY 841 TGGCGCGCTTCACGAGCGCGCGCTGCGGCTGCGGCTGCGGCTTACTGGGCGCTGAGCGCC 900
DB 957 TGGCGCGCTTCACGAGCGCGCGCTGCGGCTGCGGCTTACTGGGCGCTGAGCGCC 1016
QY 901 GAGCGCGCGCGCGCGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 960
DB 1017 GAGCGCGCGCGCGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1076
QY 961 CTGCGCGCGCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1020
DB 1077 CTGCGCGCGCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1136
QY 1021 GGGCCACTGCGCTTCGAGGCGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1080
DB 1137 GGGCCACTGCGCTTCGAGGCGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1196
QY 1081 AGCGGCTCCCGCGCGCGCGCGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1140
DB 1197 AGCGGCTCCCGCGCGCGCGCGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1256
QY 1141 GTGCGAGTGCCTGCGGCTGCGGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1200
DB 1257 GTGCGAGTGCCTGCGGCTGCGGCTGCGGCTTCTCCTCAAGCACGAGCTCAGCTCCGAG 1316
QY 1201 GGGCGGACCGGCGCGCGCGGCAAGCAAGGCGAGGTGCTGCTGCGGCGCGGCGGCGGCGGCGG 1260
DB 1317 GGGCGGACCGGCGCGCGCGGCAAGCAAGGCGAGGTGCTGCTGCGGCGCGGCGGCGGCGGCGG 1376
QY 1261 GAGATCTGCGTCTGA 1275
DB 1377 GAGATCTGCGTCTGA 1391
RESULT 2
AAS17862
ID AAS17862 standard; cDNA; 2598 BP.
XX
AC AAS17862;
XX
DT 08-MAY-2002 (first entry)
XX Human protein kinase N cDNA.
DE
KW Human; ss; protein kinase N; cytosolic; neuroprotective; cancer;
KW gene therapy; antigen; antibody; neurodegenerative disease;
KW inflammation; arteriosclerosis; psoriasis; growth disorder;
KW chromosome 16; papilloma virus infection, Alzheimer's disease.
OS Homo sapiens.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 14:17:18 ; Search time 299 Seconds
(without alignments)
11510.982 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atagagctggctgcccaga.....ccatcgagatctgcgtctga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Genesec 19Jun03.*

1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	100.0	1473	24	ABK86975 Novel human protei
2	1275	100.0	2598	24	AA517862 Human protein kina
3	1273.4	99.9	1275	22	AA506744 Polynucleotide seq
4	1273.4	99.9	1275	25	AA49419 Human serine/threo
5	1273.4	99.9	2060	24	AA34298 Human PKIN-1 cDNA
6	1273.4	99.9	3124	25	AA49420 Human serine/threo
7	1262.4	99.0	1272	25	AA49416 Human serine/threo
8	1131.8	88.8	2893	24	ABK10101 Human cDNA encodin

9	963.4	75.6	1527	21	AAA09825 Rat PK5 gene, Rat
10	931.2	73.0	936	25	AA49417 Human serine/threo
11	849.6	66.6	7301	24	AA517863 Human protein kina
12	767.6	60.2	788	24	AA520990 DNA sequence #21 r
13	767.6	60.2	788	24	AA521038 Isolated CpG islan
14	736.6	18.6	1247	24	ABK89200 DNA encoding human
15	236.6	18.6	1716	22	AA506705 Polynucleotide seq
16	235	18.4	1156	24	AA43413 Human PKIN-16 cDNA
17	196	15.4	1119	22	AA506707 Polynucleotide seq
18	196	15.4	2380	25	AA49443 Human kinase and p
19	179	14.0	1887	25	AA49413 Human serine/threo
20	147.2	11.5	1865	23	AB517453 Drosophila melanog
21	147.2	11.5	14775	23	AB517452 Drosophila melanog
22	131.4	10.3	2132	23	ABU08603 Zea mays DNA fragm
23	105.8	8.3	1435	21	AA46591 Novel protein kina
24	100.4	7.9	1788	22	AA464624 Human kinase (PKIN
25	99.8	7.8	1487	24	AA438855 Human kinase (PKIN
26	99.8	7.8	1640	22	AA08645 Human serine/threo
27	99.8	7.8	1864	24	AB076181 Human polynucleoti
28	99.8	7.8	2007	22	AA159743 Human NS cDNA sequ
29	99.6	7.8	2182	24	ABU39744 Human neuroblastom
30	98.4	7.7	774	22	AA197091 Human polynucleoti
31	98.2	7.7	1972	22	AA157962 Human Akt1 encodin
32	97.6	7.7	1443	24	ABU50837 PKB-green fluorescent
33	97.6	7.7	2181	19	AAV71037 Green fluorescent
34	97.6	7.7	2184	19	AAV71086 Human RAC protein
35	97.6	7.7	2610	18	AA64812 Human RAC protein
36	97.6	7.7	2610	18	AA67135 Human Akt-1 encodi
37	97.6	7.7	2610	20	AA241177 Human Akt-1 DNA se
38	97.6	7.7	2610	20	AA221190 Wild type human Ak
39	97.6	7.7	2610	21	AA09075 Human cDNA differe
40	97.6	7.7	2610	24	ABK84055 Human colon cancer
41	97.6	7.7	2610	24	AA28548 Human Akt2 encodin
42	96.8	7.6	718	21	AAH31053 Mouse Akt-2 cognat
43	93.4	7.3	1446	24	ABU50838 Human Akt-2 nucleo
44	93.4	7.3	1599	18	AA71252 Human Akt-2 nucleo
45	93.4	7.3	1599	21	AAA08448

ALIGNMENTS

RESULT 1
ABK86975
ID ABK86975 standard; cDNA; 1473 BP.

AC ABK86975;

XX 24-SEP-2002 (first entry)

DE Novel human protein (NHP) coding DNA.
Human; Novel human protein; NHP; gene; ss; kinase; regulatory pathway;
therapeutic; diagnostic; pharmacogenomic; antigenic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 117..1391

FT /*tag= a

FT /product= "NHP"

FT /note= "This region is specifically claimed in claim 1 of the specification"

PN WO200246428-A2.

XX 13-JUN-2002.

PD 07-DEC-2001; 2001WO-US48533.

XX 07-DEC-2000; 2000US-251941P.

XX (LEXI-) LEXICON GENETICS INC.

Db 504 AGTACGTGCCCGCGGAGATGTTCTCACACCTACGGCGATCGGAAGTTCAGTGAGC 563
Qy 446 ACACGGTGAAGCGTGTGTGACAGCTGGGCTTGGGCTTGGACTTTCATGCACGGCGGC 505
Db 564 CCCATGCCCGTTTCTACGGCGCCAGATCGTCTGACCTTTGAGTATCTGCACTCGCTGG 623
Qy 506 AGCTGGTGCACCGGACATCAAGCCCGAGACGTCCTGCTTTCGACGGGAGTCCGCC 565
Db 624 ATCTCACTACAGGAGCTGAAGCCGAGAACTGCTCAITGACAGCAGG-----GCT 677
Qy 566 GCGTAAAGCTGGCGGACTTCGGCATGACGCCCGCTGGGTGCGCGTCAAGCGCGTGA 625
Db 678 ACATTGAGTACAGACTTCGTTTCGCCAAGCGGTGAAGGGCGGCACTTGGACCTTGT 737
Qy 526 GCGGACCATCCCTACACGGCGCTGAGGTGTCCAGGCGGGCGCGCCGACGGGCTGG 685
Db 738 GCGGACCCCTGAGTACTGCGCCCTGAGATTATCTTGAGCAAAAGGCTACAAAGGCG 797
Qy 686 CGGTGGACACGGCGTGGACGT 707
Db 798 TGGACTGGTGGCCCTGGGGT 819

RESULT 15

US-09-842-307-1
; Sequence 1, Application US/09842307
; Patent No. 6562574
; GENERAL INFORMATION:
; APPLICANT: Altschuler, David M.
; APPLICANT: Hirschhorn, Joel N.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: ASSOCIATION OF PROTEIN KINASE C ZETA
; TITLE OF INVENTION: POLYMORPHISMS WITH DIABETES
; FILE REFERENCE: 2825.2011-001
; CURRENT APPLICATION NUMBER: US/09/842.307
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,468
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2146
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-307-1

Query Match 5.8%; Score 73.8; DB 4; Length 2146;
Best Local Similarity 47.8%; Pred. No. 1.2e-05;
Matches 296; Conservative 0; Mismatches 302; Indels 21; Gaps 2;
Qy 152 AGCACTAGCACTAGTCCGGAGCTGGGAGCTGGCAAGGACCTATGGGAAGGTGACCTGGTGG 211
Db 731 AGGACTTTGACCTTAATCAGAGTCATCGGGCGCGGAGCTACGCCAAGGTTCTCTCTGTC 790
Qy 212 TCTACAAAGGCGACAGGCACAAATGACCTGAAGTTGTGAACA-----GAGCA 262
Db 791 GGTGAGAGAGATGACCAAAATTACGCCATGAAGTGTGAAGAGAGCTGTGTCATG 850
Qy 263 AAACCAAGCTGAAGACTTCTACGGAGGTGAGCATCACCAAGAGCTTCTCCAGCC 322
Db 851 ATGACGAGGATATTGACTGGGTACAGACAGAGAGCAGCTTTTGGCAGGCAATCCAGCA 910
Qy 323 CCTTCATCATCAAGGTCTTTGACGTGCTTTGACAGAGGACTGTACGTCTTTGCC 382
Db 911 ACCCTTCTCTGGTCCGATTACACTCTCTCCAGACGACAAGTCGGTTGTCTCTGTCA 970
Qy 383 AGGAGTACGACCTGCTGGGACCTGTTTGATCATCATCCCTCCCGGCGGCTCCCTG 442
Db 971 TTGAGTACGTCAAGCGGGGACCTGATGTTCCATGACAGGAGAGAGAGCTCCCTG 1030
Qy 443 AGGACAGGTGAAGCGCTGTGTGAGCAGAGCTGGGCTGGGCTGAGATTTCATGACGGGC 502
Db 1031 AGGAGCAGCCAGGTTCCTACGCGCGGAGATCTGCATCGCCCTCAACTTCTCTGCACGAGA 1090

Qy 503 GGCAGCTGGTGCACCGGACATCAAGCCCGAGAACGTGCTGCTTTCGACCGGAGTGCC 562
Db 1091 GGGGGATCATCTACAGGACCTGAAGCTTGACAAACGTCTCTCTGATGCGGACGGGCACA 1150
Qy 563 GCCCGTAAAGCTGGCGGACTTCGGCATGACGGCGCGCTGGGCTGCGGCTCAAGCGC 622
Db 1151 TCAAGCTCAAGACTACGGCATGTGCAAGGAAGCCCTGGGCCCCCTGGTGACACAAACGAGCA 1210
Qy 623 TGAGCGGCACCATCCCTTTACCGCGCGCTGAGGTGTCCAGGGCGGCGCGGCCACGGC 682
Db 1211 CTTTCTGCGGAACCCCGAATTACATCGCCCCCGAAATCCTCGCGGAGAGAGTACGGGT 1270
Qy 683 TGGCGGTGGACACGGGCGTGGACGTGTGGGCTTTCGGCGTGTCTCATCTTCTGCGTGTCA 742
Db 1271 T-----CAGCGTGGACTGTGTGGCGCTGGGAGTCTCTCATGTTTGAATGATGG 1318
Qy 743 CCGGCAACTTCCCGTGGGA 761
Db 1319 CCGGGCGCTCCCGTTTCA 1337

Search completed: December 7, 2003, 16:47:45
Job time : 84 secs

Db 445 AGTACGTGCCCGGGGAGATGTTCTACACCTACGGCGGATCGGAAGGTTCAGTGAGC 504
QY 446 ACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTGGGCTGGACTTCATGACGGGGGGC 505
Db 505 CCATGCGCCGTTTCTACGCGGCCAGATGCTGACCTTTGATATCTGCACCTCGTGG 564
QY 506 AGCTGTGACCGCGGACATCAAGCCCGAGAACGCTGTGCTTTCGACCGGAGTGGCGCC 565
Db 565 ATCTCATCTACAGGACCTGAAGCGCGAGATCTGCTCAITGACACGAGG-----GCT 618
QY 566 GCGTAAAGTGGCGACTTCGGCATACGCGCGCTGGGCTGGCGGTCAAGCGGTGA 625
Db 619 ACATTCAAGTGACAGACTTCGGTTTCGCCAAGCGCGTGAAGGGCCGCACTTGGACCTTGT 678
QY 626 GCGGCACCATCCCTTACACGCGCGCTGAGGTGTGCCAAGCGCGCGCGCGCGGCTGG 685
Db 679 GCGGCACCCCTGAGTACTTGGCCCTGAGATTATCTTGAGCAAGGCTACACAGGCGG 738
QY 686 CGGTGGACACGGCGGTGACGT 707
Db 739 TGGACTGTGGGCGCTGGGGT 760

RESULT 13

US-09-394-455-5
; Sequence 5, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustin, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,771
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1133)
; NAME/KEY: misc feature
; LOCATION: (1)...(2549)
; OTHER INFORMATION: n = A,T,C or G

US-09-394-455-5
Query Match 5.8%; Score 74.4; DB 4; Length 2549;
Best Local Similarity 49.6%; Pred. No. 9.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
QY 152 AGCTACGACTAGTCCGGAGCTGGGCAAGGACCTATGGGAAGTTGACCTGTGGTGG 211
Db 205 ATCAGTTTGAACGATCAAGACCTTCGGCAAGGCTCCTTCGGGCGGTGATGCTGTGA 264
QY 212 TCTCAAGGGCAGGCAACAAATGGCACTGAAGTTTGTGAACAGAGCAAAA---CCA 268
Db 265 AACACAGGAGACCGGGAACCACTATGCGATGAAGATCTTCGACAAACAGAGGTGGTA 324
QY 269 AGCTGAAGAACTTCCTACGGAGTGGGCAAGGACCTATGGGAAGTTGACCTGTGGTGG 328
Db 325 AACTGAACAGATCGAACACACCCCTGAATGAAGGCGATCTCGAAGCTGTCACTTTC 384
QY 329 TCATCAAGGTCTTTGACGTGTCTTTTGACAGAGAGCTGTACGT---CTTTGCCGAGG 385
Db 385 CGTTCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTGTCTG 444

QY 386 AGTACGCACTGCTGGGACCTGTTTGAATCATCTCCCTCCACAGGTGGGCTCCCTGAGG 445
Db 445 AGTACGTGCCCGGGGAGATGTTCTACACCTACGGCGGATCGGAAGGTTCAGTGAGC 504
QY 446 ACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTGGGCTGGACTTCATGACGGGGGGC 505
Db 505 CCATGCGCCGTTTCTACGCGGCCAGATGCTGACCTTTGATATCTGCACCTCGTGG 564
QY 506 AGCTGTGACCGCGGACATCAAGCCCGAGAACGCTGTGCTTTCGACCGGAGTGGCGCC 565
Db 565 ATCTCATCTACAGGACCTGAAGCGCGAATCTGCTCATTTGACACGAGG-----GCT 618
QY 566 GCGTAAAGTGGCGACTTCGGCATACGCGCGCTGGGCTGGCGGTCAAGCGGTGA 625
Db 619 ACATTCAAGTGACAGACTTCGGTTTCGCCAAGCGCGTGAAGGGCCGCACTTGGACCTTGT 678
QY 626 GCGGCACCATCCCTTACACGCGCGCTGAGGTGTGCCAAGCGCGCGCGCGCGGCTGG 685
Db 679 GCGGCACCCCTGAGTACTTGGCCCTGAGATTATCTTGAGCAAGGCTACACAGGCGG 738
QY 686 CGGTGGACACGGCGGTGACGT 707
Db 739 TGGACTGTGGGCGCTGGGGT 760

RESULT 14

US-09-394-455-35
; Sequence 35, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustin, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,771
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 2608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)...(1192)
; OTHER INFORMATION: n = A,T,C or G

US-09-394-455-35
Query Match 5.8%; Score 74.4; DB 4; Length 2608;
Best Local Similarity 49.6%; Pred. No. 9.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
QY 152 AGCTACGACTAGTCCGGAGCTGGGCAAGGACCTATGGGAAGTTGACCTGTGGTGG 211
Db 264 ATCAGTTTGAACGATCAAGACCTTCGGCAAGGCTCCTTCGGGCGGTGATGCTGTGA 323
QY 212 TCTCAAGGGCAGGCAACAAATGGCACTGAAGTTTGTGAACAGAGCAAAA---CCA 268
Db 324 AACACAGGAGACCGGGAACCACTATGCGATGAAGATCTTCGACAAACAGAGGTGGTA 383
QY 269 AGCTGAAGAACTTCCTACGGAGTGGGCAAGGACCTATGGGAAGTTGACCTGTGGTGG 328
Db 384 AACTGAACAGATCGAACACACCCCTGAATGAAGGCGCATCTCGAAGCTGTCACTTTC 443
QY 329 TCATCAAGGTCTTTGACGTGTCTTTTGACAGAGAGCTGTACGT---CTTTGCCGAGG 385
Db 444 CGTTCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTGTATGG 503
QY 386 AGTACGCACTGCTGGGACCTGTTTGAATCATCTCCCTCCACAGGTGGGCTCCCTGAGG 445

Db 124 AGRAGAATGATACCAAGAGATGCGCAATGAAGTATACATGAATAAACAAGTGGG 183
Qy 272 TGAAGAACTTCTTACGGGAGGTGAGCA-----TCACCAACAGCCTCTCTCCAGCCCT 325
Db 184 AGCGCAATGAAGTGAAGAAATGTCTTCAAGGAATCCAGATCATCGAGGCTCGAGCACC 243
Qy 326 TCATCATCAAGTCTTTGACGTGCTTTGAGACAGAGGACTGCTAGTCTTTGCCCAGG 385
Db 244 CTTTCTCGTTAAATTTGTGGTATCTCTTCCAAATGAGGAAGACATGTTCAATGTTGG 303
Qy 386 AGTACGACCTGCTGGGACCTGTTGACATCATCCCTCCCGAGTGGGCTCCCTGAGG 445
Db 304 ACCTCTCTGGTGGAGACCTGCTATACCTGCAACAGAGCTCCACTTCAAGNAG 363
Qy 446 ACAGGTGAAGCGTGTGTGAGCAGCTGGGCTGGGCTGGACTTTCATGACGGGGGG 505
Db 364 AAACAGTGAAGCTTCTATCTGTGAGTGTCTATGGGCTGGACTGTCAGAACACAGC 423
Qy 506 AGCTGGTGCACCGGACATCAAGCCCGAGACGCTGCTCTTGCACGGGAGTGGCCGC 565
Db 424 GCATCATTCACAGGATATGAAGCTGACAAATATTTACTTGAACACATGGGCACTGC 483
Qy 566 GCGTAAAGCTGGCCGACTTCGGCATGACGCGCGGCTGGGCTGGCGGTCAAGGCGTGA 625
Db 484 ACATCAGAGATTCAACATTCTCGATGCTGCCAGGAGACACAGATTACCACTGG 543
Qy 626 GCGCACATCCCTTACAGCGCGCTGAGGTGTC 660
Db 544 CTGCGCAACAGCCTTACATGGCCTGAGATGTTTC 578

RESULT 11
US-09-394-455-3
; Sequence 3, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustín, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; PRIOR FILING DATE: 1999-09-10
; PRIOR FILING DATE: 1998-09-10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1008)
US-09-394-455-3

Query Match 5.8%; Score 74.4; DB 4; Length 1008;
Best Local Similarity 49.6%; Pred. No. 8.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
Qy 152 AGCACTAGCACTAGTCCGGAGTGGCAAGGACCTATGGAGGTTGACCTGGTGG 211
Db 80 ATCAGTTTGAACGAATCAAGACCTTCGGCAGCGGCTCTCTGGGCGGTGATGTTGGTGA 139
Qy 212 TCTACAGGCGACAGGCACAAAATGGCACTGAAGTTTGTGAACAAGACAAA---CCA 268
Db 140 AACCAAGGAGACCGGGAACCACTATGCCATGAAGATCTCTGACNAACAGAGGTTGTA 199
Qy 269 AGCTGAAGAACTTCTTACGGGAGGTGAGCATCACCAAGCCTCTCTCCAGCCCTTCA 328
Db 200 AACTGAACACAGATGAAACACACCTGTAATGAAGAGCGCATCTCTGCAAGCTGTCAACTTC 259

Qy 329 TCATCAAGGTCTTTGACGTGCTTTTGAGACAGAGGACTGCTAGCT---CTTTGCCCAGG 385
Db 260 CGTTCCTCGTCAAACTCGAGTTCTCTCTCAAGGACAACTCAAACTTATACATGTCATGG 319
Qy 386 AGTACGACCTGCTGGGACCTGTTTGAATCATCCCTCCCGAGTGGGCTCCCTGAGG 445
Db 320 AGTACGTGCGCGGGGAGATGTTCTCACATACGCGGATCGGAAGGTTAGTGAGC 379
Qy 446 ACAGGTGAAGCGTGTGTGACAGCTGGGCTTGGGCTGGACTTCATGACAGGGGGC 505
Db 380 CCATGCGCCCTTCTACGCGCCAGATCGTCTGACCTTTGAGTATCTGCACTCGCTGG 439
Qy 506 AGCTGTGACCGGACATCAAGCCCGAGAGCTGCTGTGTCGACCGGAGTGGCCGC 565
Db 440 ATCTCATCTTACAGGACCTGAAGCGGAGAACTCTGCTCATTTGACCAGCAGG------GCT 493
Qy 566 GCGTAAAGCTGGCGACTTCGGCATGACGCGCGGCTGGGCTGGCGGTCAAGGCGCTGA 625
Db 494 ACATTCAGGTGACAGACTTCGTTTCGCAAGCGGTGAAGGGCGGCACTTGGACCTGT 553
Qy 626 GCGGACCATCCCTTACAGCGGCGCTGAGGTGTCAGGCGGCGCGCGGCGGCTGG 685
Db 554 GCGGACCCCTGAGTACCTGCGCCCTGAGATTATCTCTGAGCAAGGCTACAAACAAGCCG 613
Qy 686 CGGTGGACAGCGGCGTGGACGT 707
Db 614 TGGACTGGTGGGCGCTGGGGT 635

RESULT 12
US-09-467-082-3
; Sequence 3, Application US/09467082
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF PKA CATALYTIC SUBUNIT C-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0088
; CURRENT APPLICATION NUMBER: US/09/467,082
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(1136)
; NAME/KEY: unsure
; LOCATION: 6
; OTHER INFORMATION: unknown
US-09-467-082-3

Query Match 5.8%; Score 74.4; DB 3; Length 2549;
Best Local Similarity 49.6%; Pred. No. 9.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
Qy 152 AGCACTAGCACTAGTCCGGAGTGGGCAAGGACCTATGGAGGTTGACCTGGTGG 211
Db 205 ATCAGTTTGAACGAATCAAGACCTTCGGCAGCGGCTCTCTGGGCGGTGATGTTGGTGA 264
Qy 212 TCTACAGGCGACAGGCACAAAATGGCACTGAAGTTTGTGAACAAGACAAA---CCA 268
Db 265 AACCAAGGAGACCGGGAACCACTATGCCATGAAGATCTCTGCAACAAGAGGTTGTA 324
Qy 269 AGCTGAAGAACTTCTTACGGGAGTGAATCAACAGCCTCTCTCCAGCCCTTCA 328
Db 325 AACTGAAACAGATCAACACACCTGATGAAGAGCCATCTCTGCAAGCTGTCAACTTC 384
Qy 329 TCATCAAGGTCTTTGACGTGCTTTTGAGACAGAGGACTGCTAGCT---CTTTGCCCAGG 385
Db 385 CGTTCCTCGTCAAACTCGAGTTCTCTCTCAAGGACAACTCAAACTTATACATGTCATGG 444
Qy 386 AGTACGACCTGCTGGGAGCTGTTTGAATCATCCCTCCCGAGTGGGCTCCCTGAGG 445

```
Db 540 CTTTCGACTATCTCAAACTCCTTGGCAAGGAAACCTTTGGCAAAAGTCATCTGTGGCGGA 599
Qy 216 CAAGGCGACAGGACAAATAATGGCACTGAAGTTTGTGAAACAAGAGCAAAACCAAGCTGAA 275
Db 600 GAAGGCGACAGGCGGCTACTACGCCATGAAGATCCTCGAAAGGAAGTATCATTTGCCAA 659
Qy 276 GAATCTTCTACGGAGGTGAGTACACCAAG-----CCTTCCTCCAGCCCTTCAT 329
Db 660 GGATGAAGTTCGCTCACACAGTACCGAGAGCGGGTCTCTCAGAAACACCAAGGACCCCGTT 719
Qy 330 CATCAAGGTCTTTGACGTGTCCTTTGAGACAGAGACTGCTACTCTTTGCCAGAGTA 389
Db 720 CCTCACTGCGTGAAGTATGCTTCCAGACCCAGCCGCTGTGCTTTGATGAGTA 779
Qy 390 CGCACCTGCTGGGACCTGTTGACATCATCCCTCCACAGGTGGGCTCCCTGAGGACAC 449
Db 780 TGCCAAAGGGGTGAGCTGTTCTTCCACCTGTCCGGGAGCGTCTTTCACAGAGAGCG 839
Qy 450 GGTGAAGCGTGTGTGACAGAGCTGGGCTTGGCTGTGACTTATGACAGCGGCGGAGCT 509
Db 840 GGCCCGGTTTATGGTGACAGATTTGCTCGGCTCTTGAATGCTGCACTCGCGGAGCT 899
Qy 510 GGTCCAGCGACATCAAGCCCGAGAGAGTGTCTGTTCACCGCGAGTCCCGCGGT 569
Db 900 GGTATACCGGACATCAAGCTGGAAACCTCATGCTGGACAAAGATGGCCACATCAAGAT 959
Qy 570 AAAGCTGGCCGACTTCGCGATGACCGCGCGGTGGGCTGCGCGTCAAGCGCGTGAAGCGG 629
Db 960 CACTGACTTTGGGCTCTGCAAGAGGACATCAGTGACGGGCGCCACCATGAAACCTTCTG 1019
Qy 630 CACATCCTTACAGGCGCTGAGGTGTGCGAGCGGCGCGCGCGCGCGCGGCTGGCGGT 689
Db 1020 TGGACCGCGGAGTACCTTGGCGGCTGAGGTGCTGAGGACATGACTATGGCGGCGCGT 1079
Qy 690 GGACAGCGGCTGAGCGTGTGGGCTTGGGCTGCTCATCTTCTGCGTCAAGCGTCAAGCGG 749
Db 1080 GGACTGTGGGCGTGGTGTGGTGTGATGATGACAGATGATGCGCGCGCGCTGGCCCTCTA 1139
Qy 750 CTTCCCTGGAGGCGGCGTGGGCGCGCGAGCGCTTCTTCAGAGAGTTC 798
Db 1140 CAACGAGGACACGAGCGCTTCTGAGCTCATCTCATGAGAGATC 1188
```

```
RESULT 9
US-09-428-711A-15
; Sequence 15, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shirasawa, Takuji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: No. 6358720uchi, Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP98/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 5228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)...(3418)
US-09-428-711A-15
Query Match 6.4%; Score 82.2; DB 4; Length 5228;
Best Local Similarity 46.4%; Pred. No. 3.6e-07;
```

```
Matches 267; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
Qy 236 TGGCACTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGA 295
Db 396 TCGCCGTCAGTGCATTAACAAGAGAACTCTGCCAAGTCTCAGACGCTGCTGGGAAGG 455
Qy 296 GATCAACCAAGACCTCTCTCTCAGCGCTTATCATCAAGGTCTTTGAGTGGTCTTTG 355
Db 456 AAATCAAAATCTTGAAGAACTGAAACAAGAAACATCTGTGGCCCTGTACGACTTCCAGG 515
Qy 356 AGACAGAGGACTCTAGCTTTTGGCCCAAGAGTACGACCTGCTGGGAGCTTTTGACA 415
Db 516 AATATGGCTAATTTCTGCTTACCTGTTATGGAGTACTGCAACGGTGGGAGCTTGGCCACT 575
Qy 416 TCATCTCCTCTCCCAAGTGGGCTCCTCTGAGACACGGTGAAGCGCTGTGTGACAGCTGG 475
Db 576 ACCTGCACGCCATGCGCAGCTGAGCGAGGACACCATCAGGCTCTTCTCTGACAGATCG 635
Qy 476 GCTTGGCGCTGGAGCTTATGACAGGGGGGAGCTGTGTGACACCGGACATCAAGCCGAGA 535
Db 636 CGGCGGCGCATGCGGCTTCTGCACAGCAAGGCAATCATCCACCGCGACTGAAACCCGAGA 695
Qy 536 ACCTGCTGTCTGTCGACCGGAGTGGCGCGCTAAAGCTGGCGACTTCGGCATGACGC 595
Db 696 ACATCTGCTGTCTCAACCCCGCGCGCGCGCGCCCAACCCCAACAGCATCCGCGTCAAGA 755
Qy 596 GCGCGCTGGGCTGCGCGCTCAAGCGGTGAGCGGCGACCATCCCTTACA CGGCGCTGAGG 655
Db 756 TCGCTGACTTCGGCTTCGCGCGTACTCTCCAGAGCAACATGATGGCGGCCACACTCTGCG 815
Qy 656 TGTGCCAGGCGCGCGCGCGCGAGCGGCTGGCGGTGACACGCGCGGTGGAGCTGTGGCCT 715
Db 816 GCTCCCGCATGTATGCGCGCGCGCGCGCTCATCATGTGCCAGCACTACGACGGAAGCGG 875
Qy 716 TCGCGGTGCTCATCTTCTGCGTCTCACCGCAACTTCCCGTGGGAGCGCGCTCGGGCG 775
Db 876 ACCTGTGAGCATCGGACCATGCTCTACAGTGCCTGACGCGGAGGCGCCCTTCCAGG 935
Qy 776 CCGAGCGCTTCTTTCGAGAGTTCGTGCGTGGCGAG 810
Db 936 CCAGCAGCGCCCGAGGACCTCGCGCTGTCTACAG 970
```

```
RESULT 10
US-09-801-876B-1
; Sequence 1, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human
US-09-801-876B-1
Query Match 5.9%; Score 75.4; DB 4; Length 1485;
Best Local Similarity 48.2%; Pred. No. 5.8e-06;
Matches 248; Conservative 0; Mismatches 261; Indels 6; Gaps 1;
```

```
Qy 152 AGCACTAGCACTAGTCCGGAGCTGGCAAGGACCTATGGGAAGGTTGACCTGTGG 211
Db 64 ACCACTTTGAAATTTTCGAGGCCATTGGGAAGGCACTTTGGGAGGCTGTGCAATTGAC 123
Qy 212 TCTCAAGGCGACAGGCACAAAATGCACTGAAGTTTGTGAACAGAGCAAAACCAAGC 271
```

QY 609 CCGGCTCAAGCGGTGAGCGGCACATCCCTTACACGGCGCTGAGGTGTGCCAGGCGG 669
Db 1107 TGCACCATGAAGACCTTTTTCGGGCACACCTGAGTACCTGCGCCCGAGGTGCTGGAGGA 1166
QY 669 CCGGCGCAGCGGTGGCGGTGGACACGGGGTGGAGTGTGGGCCCTTGGCGTGTCTCAT 728
Db 1167 CAATGACTACGCGCGTGCAGTGGAGTGTGGGGCTGGCGTGTCTCATGTACGAGATGAT 1226
QY 729 CTTCTGCGTGTCTACCGCGCAACTTCCCGTGGGAGCGCGCTGCGGCGCGCAGCGCTTCTT 788
Db 1227 GTGCGTGTCTGCGCTTCTTACACACGAGGACCATGAGAAGCTTTTGTAGCTCATCTCAT 1286
QY 789 CGAGGAGTTC 798
Db 1287 GGAGGAGATC 1296

RESULT 7
US-09-256-465-1
; Sequence 1, Application US/09256465
; Patent No. 6043090
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowse
; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-2 EXPRESSION
; FILE REFERENCE: RTS-0035
; CURRENT APPLICATION NUMBER: US/09/256.465
; CURRENT FILING DATE: 1999-02-23
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(1533)
US-09-256-465-1

Query Match
Best Local Similarity 7.3%; Score 93.4; DB 3; Length 1599;
Mismatches 332; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 96 CACTGAAGACATGACGCGCTGACTCTCCGACACTGCGCCCGCAGCGACGTCAACAGCA 155
Db 480 CACGACTGAGGAGATGGAAGTGGCGGTGACGAGCGACCGGCTAAAGTGACCATGAATGA 539
QY 156 CTAGCAACTAGTCCGGGAGCTGGGCAAGGACCTATGGGAAGTGTGACCTGTGGTCTA 215
Db 540 CTTGCACTATCTCAAACTCTTGGCAAGGAACTTTGGCAAGTCACTCTGTGGTGGGA 599
QY 216 CAAGGCAACAGGCAAAATGACCTGAAGTTGTGAACAGAGCAAAACCAAGCTGAA 275
Db 600 GAAGGCCACTGGCGCTACTACGCCATGAACATCTGCGAAAGGAAGTCAATTTGCCAA 659
QY 276 GAATCTCTACGGGAGTGACATCAACAG-----CCTCTCTCCAGCCCTTCAT 329
Db 660 GGATGAAGTCTCTACAGCTCACAGTCCAGAGCCGGGTCTCCAGAACACAGGACCCCGTT 719
QY 330 CATCAAGTCTTTGACGTGGTCTTTGACAGAGAGTGTACCTTTTGGCCAGGAGTA 389
Db 720 CTTCACTGCGTGAAGTATGCTTCCAGACCCAGACCGCTGTGCTTTGTGATGGAGTA 779
QY 390 CGCACCTCTGGGACCTGTTTGACATCATCCCTCCCGAGGTGGGCTCCCTGTAGGACAC 449
Db 780 TGCCAAACGGGGTGTGAGTGTCTTCCACCTCTCCCGGAGCGTGTCTTACAGAGGAGCG 839
QY 450 GGTGAAGCGCTGTGTGACGAGCTGGGCTGGCTGGCTGACATTCATGACGCGCGGAGCT 509
Db 840 GGCCCGGTTTATGGTGACAGATGTCTCGGCTCTTGAGTACTTGTGACTCGCGGAGCT 899
QY 510 GGTGACCGGACATCAAGCCCGAGAGCGTGTCTGTCTTTCACCGGAGTGTCCCGCGT 569
Db 900 GGTATACCGGACATCAAGCTGGAAACCTCATGCTGGACAAAGATGGCCACATCAAGAT 959

RESULT 8

US-09-167-322-3
; Sequence 3, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/167,322
FILING DATE: 07-Oct-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/00582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 7933-33 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-167-322-3

Query Match 7.3%; Score 95.4; DB 4; Length 1599;
Best Local Similarity 46.8%; Pred. No. 2.5e-09;
Matches 332; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 96 CACTGAAGACATGACGCGCTGACTCTCCGACACTGCGCCCGCAGCGACGTCAACAGCA 155
Db 480 CACGACTGAGGAGATGGAAGTGGCGGTGACGAGCGACCGGCTAAAGTGACCATGAATGA 539
QY 156 CTAGCAACTAGTCCGGGAGCTGGGCAAGGACCTATGGGAAGTGTGACCTGTGGTCTA 215

QY 729 CTTCTCGTCTCACCGGCAACTTCCGTCGGAGCGCGCTCGGGCCGCCGACGCTTCTT 788
Db 1770 GTGGTTCGCTGCCCTTCTACACAGGACCATGAGAAGCTTTTGTAGCTCATCTCTCAT 1829
QY 789 CGAGGAGTTC 798
Db 1830 GGAGGAGATC 1839

RESULT 5

US-09-212-771-1
; Sequence 1, Application US/09212771
; Patent No. 5958773
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-1 EXPRESSION
; FILE REFERENCE: RTS-0034
; CURRENT APPLICATION NUMBER: US/09/212.771
; CURRENT FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1641)
US-09-212-771-1

Query Match 7.7%; Score 97.6; DB 2; Length 2610;
Best Local Similarity 48.8%; Pred. No. 4.3e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

QY 138 CAGCGAGCTACCCAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 627 CCGGTGACCATGAACGAGTTTGAGTACCTGAGAGCTGCTGGCAAGGCACTTTCCGCAA 686
QY 198 GGTTCAGCTGCTGTCTACAAAGGCGACAGGCAAAAATGCGCACTGAAGTTTGTGAACA 257
Db 587 GGTGATCTGTGAGGAGAGGCGCAAGCGCTACTACGCCATGAGATCCTCAAGAA 746
QY 258 GAGCAAAACCAAGCTGAAGAACTTCCCTACGGGAGGTGAGCATCACCAACAGCCTCTC 314
Db 747 GGAAGTCATCTGTGCGCAAGGACGAGTGGCCACACACTCACCGAGAACCGCGCTCTGCA 806
QY 315 ---CTCCAGCCCTTCATCATCAAGTCTTTGAGCTGTCTTTGAGACAGAGGACTGCTA 371
Db 807 GAACTCCAGGACCCCTTCTCAGCCCTTGAAGTACTCTTTCCAGACCCACGACCGCCT 866
QY 372 CGTCTTTGCCAGGAGTACGACCTGCTGGGAGCTCTTTGAGCATCATCCTCCCGAGGT 431
Db 867 CTGCTTTGTATGAGTACCGCAACGGGGGCGAGCTGTTCTTCCACTGTCCGGGAACG 926
QY 432 GGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTGGGACTT 491
Db 927 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 986
QY 492 CATGCAC---GGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACCTGTGCTGTT 548
Db 987 CTTGCACTCGGAGAAAGACGTGGTGTACCGGACCTCAAGCTGGAGAACCTCATGCTGGA 1046
QY 549 CGACCGGAGTCCCGCGCTAAAGCTGGCGGACCTTCGGCATGACCGCCCGCGTGGGCTG 608
Db 1047 CAAGGACGGCACATTAAGATCACAGACTTCGGGCTGTGCAAGGAGGGGATCAAGACGG 1106
QY 609 CCGGCTTAAGCGCGTGTGAGGCGCACATCCCTTACACGGCGCTGAGTGTGCGAGCGGG 668
Db 1107 TGCCACCATGAAGACCTTTTGGGCGACACCTGAGTACCTGGGCCCGGAGGTGCTGGAGGA 1166
QY 669 CCGGCGGAGCGGTGCGGTGGACACGCGGCTGAGACCTGTGGGCTTCGGGCTGCTCAT 728

Db 1167 CAATGACTAGGCGCGTSCAGTGGACTGCTGGGCGCTGGGCGTGTGATGTACGAGATGAT 1226
QY 729 CTTCTCGTCTCACCGGCAACTTCCGTCGGAGCGCGCTCGGGCCGCCGACGCTTCTT 788
Db 1227 GTCGGTGCGCTGCCCTTCTACAAACGAGGACCATGAGAAGCTTTTGTAGCTCATCTCAT 1286
QY 789 CGAGGAGTTC 798
Db 1287 GGAGGAGATC 1296

RESULT 6

US-09-091-058-1
; Sequence 1, Application US/09091058
; Patent No. 6054285
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Frech, Matthias
; TITLE OF INVENTION: Screening Method
; FILE REFERENCE: 4-20683/A/20684/PCT
; CURRENT APPLICATION NUMBER: US/09/091.058
; CURRENT FILING DATE: 1998-06-10
; EARLIER APPLICATION NUMBER: PCT/EP96/04814
; EARLIER FILING DATE: 1996-11-05
; EARLIER APPLICATION NUMBER: 9525703.6
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1641)
US-09-091-058-1

Query Match 7.7%; Score 97.6; DB 3; Length 2610;
Best Local Similarity 48.8%; Pred. No. 4.3e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

QY 138 CAGCGAGCTACCAAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 627 CCGGTGACCATGAACGAGTTTGAGTACCTGAGAGCTGCTGGCAAGGCACTTTCCGCAA 686
QY 198 GGTTCAGCTGCTGTCTACAAAGGCGACAGGCAAAAATGCGCACTGAAGTTTGTGAACA 257
Db 587 GGTGATCTGTGAGGAGAGGCGCAAGCGCTACTACGCCATGAGATCCTCAAGAA 746
QY 258 GAGCAAAACCAAGCTGAAGAACTTCCCTACGGGAGGTGAGCATCACCAACAGCCTCTC 314
Db 747 GGAAGTCATCTGTGCGCAAGGACGAGTGGCCACACACTCACCGAGAACCGCGCTCTGCA 806
QY 315 ---CTCCAGCCCTTCTATCATCAAGTCTTTGAGCTGTCTTTGAGACAGAGGACTGCTA 371
Db 807 GAACTCCAGGACCCCTTCTCAGCCCTTGAAGTACTCTTTCCAGACCCACGACCGCCT 866
QY 372 CGTCTTTGCCAGGAGTACGACCTGCTGGGAGCTCTTTGAGCATCATCCTCCCGAGGT 431
Db 867 CTGCTTTGTATGAGTACCGCAACGGGGGCGAGCTGTTCTTCCACTGTCCGGGAACG 926
QY 432 GGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTGGGACTT 491
Db 927 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 986
QY 492 CATGCAC---GGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACCTGTGCTGTT 548
Db 987 CTTGCACTCGGAGAAAGACGTGGTGTACCGGACCTCAAGCTGGAGAACCTCATGCTGGA 1046
QY 549 CGACCGGAGTCCCGCGCTAAAGCTGGCGGACCTTCGGCATGACCGCCCGCGTGGGCTG 608
Db 1047 CAAGGACGGCACATTAAGATCACAGACTTCGGGCTGTGCAAGGAGGGGATCAAGACGG 1106

RESULT 3
US-09-417-197-70
; Sequence 70, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 70
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PKB-EGFP fusion
; NAME/KEY: CDS
; LOCATION: (1)..(2178)
US-09-417-197-70

Query Match 7.7%; Score 97.6; DB 4; Length 2181;
Best Local Similarity 48.8%; Pred. No. 4.2e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;
QY 138 CAGCGAGTCCACCAAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 429 CCGGCTGACCAATGAACGAGTTGAGTACCTGAAGCTGCTGGCAAGGCACTTTCCGCAA 488
QY 198 GGTTCACCTGTGTCTACAAAGGCGACAGGACAAATAATGCACTGAAGTTTGTGAACAA 257
Db 489 GGTGATCTGTGTGAAGGAGGAGGCAAGGCGCTACTACGCCATGAAGATCCTCAAGAA 548
QY 258 GAGCAAAACCAAGTGAAGAACTTCTCAAGGAGGTGAGCATACCAACAGCCTCTC--- 314
Db 549 GGAAGTCATCTGTGCAAGGAGGAGGTGGCCCAACACACTACCGAGAACCGCGTCTGCA 608
QY 315 ---CTCCAGCCCTTCATCATCAGGTCTTTGAGTGTCTTTGAGACAGAGACTGCTA 371
Db 609 GAATCCAGGACCCCTTCTCTCAGCCCTGAAGTACTCTTTCCAGACCCACAGCCGCT 668
QY 372 CGTCTTTGCCAGGACGACGACCTGTCTGGGACCTGTTGACATCATCCCTCCCGAGGT 431
Db 669 CTGCTTTGTATGAGTACGACCAAGCGGGGCGAGCTGTCTTCCACCTGTCCCGGACG 728
QY 432 GGGGCTCCCTGAGGACAGCGTGAAGCGCTGTGTGACAGAGTGGGCTTGGCGTGGACTT 491
Db 729 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTACGCCCTGGACTA 788
QY 492 CATGCAC---GGGCGGAGCTGTGTGACCGGACATCAGCCGCACTACAGCCGAGTGTCTGTT 548
Db 789 CCTGCATTCGAGAAGAACGCTGTGTACCGGACCTCAAGCTGAGAACCTCATGCTGGA 848
QY 549 CGACCGGAGTGGCGCCCGCTAAAGCTGGCCGACTTCGGCATGACGCGCCCGCTGGGCTG 608
Db 849 CAAGGACGGGCACATTAGATCAAGACTTCGGGCTGTGCAAGGAGGGGATCAAGGACGG 908
QY 609 CCGGCTCAAGCGGTGAGCGGACCATCCCTTTACACGGCGCTTGAAGTGTGCCAGGCGGG 668
Db 909 TGCACCATGAAGACCTTTTGGCGCACACCTGAGTACCTGCGCCCGAGGTGCTGGAGGA 968
QY 669 CCGGCGGACGGCTGGCGGTGAGACAGCGGCTGACGCTGTGGCCCTTCGCGCTGCTCAT 728
Db 969 CAATGACTACGCGGCTGCAGTGGACTGTGTGGGGCTGGGGCTGTGATGTACGAGATGAT 1028
QY 729 CTTCTGCTGTCTACCGGCAACTTCCCTGGGAGGCGCGCTCGGCGCGCGCTTCTT 788
Db 1029 GTGGGTGCTGCTTCTTCAACACGAGGACATGAGAAGCTTTTGTAGCTCATCTCAT 1088

QY 789 CGAGGAGTTC 798
Db 1089 GGAGGAGATC 1098

RESULT 4

US-09-417-197-138
; Sequence 138, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 138
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PKB fusion
; NAME/KEY: CDS
; LOCATION: (1)..(2181)
US-09-417-197-138

Query Match 7.7%; Score 97.6; DB 4; Length 2184;
Best Local Similarity 48.8%; Pred. No. 4.2e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;
QY 138 CAGCGAGTCCACCAAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 1170 CCGGCTGACCAATGAACGAGTTGAGTACCTGAAGCTGCTGGCAAGGCACTTTCCGCAA 1229
QY 198 GGTTCACCTGTGTCTACAAAGGCGACAGGACAAATAATGCACTGAAGTTTGTGAACAA 257
Db 1230 GGTGATCTGTGTGAAGGAGGAGGCAAGGCGCTACTACGCCATGAAGATCCTCAAGAA 1289
QY 258 GAGCAAAACCAAGTGAAGAACTTCTCAAGGAGGTGAGCATACCAACAGCCTCTC--- 314
Db 1290 GGAAGTCATCTGTGCGCAAGGAGGAGGTGGCCCAACACACTACCGAGAACCGCGTCTGCA 1349
QY 315 ---CTCCAGCCCTTCATCATCAGGTCTTTGAGTGTCTTTGAGACAGAGACTGCTA 371
Db 1350 GAATCCAGGACCCCTTCTCTCAGCCCTGAAGTACTCTTTCCAGACCCACAGCCGCT 1409
QY 372 CGTCTTTGCCAGGAGTACGACCTGTCTGGGACCTGTTTGAATCATTCCTCCCGAGGT 431
Db 1410 CTGCTTTGTATGAGTACGCCAAGCGGGGCGAGCTGTCTTCCACCTGTCCCGGGAACG 1469
QY 432 GGGGCTCCCTGAGGACAGCGTGAAGCGCTGTGTGACAGAGTGGGCTTGGCGTGGACTT 491
Db 1470 TGTGTTCTCCGAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTACGCCCTGGA 1529
QY 492 CATGCAC---GGGCGGAGCTGTGTGACCGGACATCAAGCCGAGAACGCTGTGCTGTT 548
Db 1530 CTTGCACTCGAGAGAGAACGTGTGTACCGGACCTCAAGCTGAGAACCTCATGCTGGA 1589
QY 549 CGACCGGAGTGGCGCGCTAAAGCTGCGGACTTGGCATGACGCGCCCGCTGGGCTG 608
Db 1590 CAAGGACGGGCACATTAAATGATCAGACTTCGGGCTGTGCAAGGAGGGATCAAGGACGG 1649
QY 609 CCGGCTCAAGCGGTGAGCGGACCATCCCTTTACAGCGCGCTTGAAGTGTGCCAGGCGGG 668
Db 1650 TGCACCATGAAGACCTTTTTCGGGCAACCTGAGTACTTGGCCCCCGAGGTGTGGAGGA 1709
QY 669 CCGGCGGAGGCTGGCGGTGGAGACGCGGCTGGAGTGTGGGCTTTCGCGCTGCTCAT 728
Db 1710 CAATGACTACGCGGCTGACGTGGACTGTGGGCGCTGGGCGTGTGATGTACGAGATGAT 1769

Db 1134 GAGGACTGTACGTCCTTTGCCCAGGAGTACGCACCTGCTGGGACCTGTTTGATCATCATC 1193
QY 421 CCTCCCCAGGTGGGCTCCCTGAGGACACCGGTGAAGCGCTGTGTGACGACGTGGGCGTG 480
Db 1194 CCTCCCCAGGTGGGCTCCCTGAGGACACCGGTGAAGCGCTGTGTGACGACGTGGGCGTG 1253
QY 481 GCGCTGAGCTTCATGACCGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAGAGTG 540
Db 1254 GCGCTGAGCTTCATGACCGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAGAGTG 1313
QY 541 CTGCTGTTCACCGCGAGTGCAGCGGCTGAAGCTGACCGGCTGACGACATCAAGCCCGAGAGAGTG 600
Db 1314 CTGCTGTTCACCGCGAGTGCAGCGGCTGAAGCTGACCGGCTGACGACATCAAGCCCGAGAGAGTG 1373
QY 601 GTGGGCTGCCCGCTGAAGCGGCTGAGCGGACACATCCCTTACACCGGCGCTGAGGTGTGC 660
Db 1374 GTGGGCTGCCCGCTGAAGCGGCTGAGCGGACACATCCCTTACACCGGCGCTGAGGTGTGC 1433
QY 661 CAGCGCGGCGCGCGCGGCTGCGGTGGAACACCGGCGGTGACGTGTGGGCTTTCGCG 720
Db 1434 CAGCGCGGCGCGCGCGGCTGCGGTGGAACACCGGCGGTGACGTGTGGGCTTTCGCG 1493
QY 721 GTGCTCATCTTCGCTGCTCACCGGCACTTCCGCTGGGAGGCGGCTGCGGCGCGGAC 780
Db 1494 GTGCTCATCTTCGCTGCTCACCGGCACTTCCGCTGGGAGGCGGCTGCGGCGCGGAC 1553
QY 781 GCCTTCTTCGAGGAGTTCGTGCGCTGCGAGCGGCGCGCTTCCGCGGCGCTTTCGAG 840
Db 1554 GCCTTCTTCGAGGAGTTCGTGCGCTGCGAGCGGCGCGCTTCCGCGGCGCTTTCGAG 1613
QY 841 TGGCGCGCTTTCACGAGCGCGCGCTGCGGATGTTCCAGCGCTTACTTGGGCGCTGGAGCC 900
Db 1614 TGGCGCGCTTTCACGAGCGCGCGCTGCGGATGTTCCAGCGCTTACTTGGGCGCTGGAGCC 1673
QY 901 GAGCGCGCGCGCGCGCGGAGGCTTCCGCTTCCCTCAAGCAGGAGCTCACGCTCCGAG 960
Db 1674 GAGCGCGCGCGCGCGCGGAGGAGTTCGCTTCCCTCAAGCAGGAGCTCACGCTCCGAG 1733
QY 961 CTGCGCGCGCGCGCGCTGCGACCGCGCGGCAAGCCCGCGGCGGACCGCGCGCGCGCG 1020
Db 1734 CTGCGCGCGCGCGCGCTGCGACCGCGCGGCGGCAAGCCCGCGGCGGACCGCGCGCGCG 1793
QY 1021 GGGCGACTGCGCTGAGGCGCGCTGGGCGCTCAAGCGGAGCGGTGTGACCGAGAGCGG 1080
Db 1794 GGGCGACTGCGCTGAGGCGCGCTGGGCGCTCAAGCGGAGCGGTGTGACCGAGAGCGG 1853
QY 1081 AGCGGCTCCCGCGCGCGCGCGCGCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1140
Db 1854 AGCGGCTCCCGCGCGCGCGCGCGCGCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1913
QY 1141 GTGCCAGTGCCTGCGCGTGCCTGTGCGCGAGCGCGCGCTAGTCCCGCGAGGCGCGCGCG 1200
Db 1914 GTGCCAGTGCCTGCGCGTGCCTGTGCGCGAGCGCGCGCTAGTCCCGCGAGGCGCGCGCG 1973
QY 1201 GCGCGGACCGGCGCGCGCGGACGAGGAGCAAGGCGAGGTGTGCTGGCCACCGCGCATC 1260
Db 1974 GCGCGGACCGGCGCGCGCGGACGAGGAGCAAGGCGAGGTGTGCTGGCCACCGCGCATC 2033
QY 1261 GAGATCTGCGTCTGA 1275
Db 2034 GAGATCTGCGTCTGA 2048

RESULT 2

US-09-816-094-3
; Sequence 3, Application US/09816094
; Patent No. 6534299
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP

; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Human
US-09-816-094-3

Query Match 66.6%; Score 849.6; DB 4; Length 7301;
Best Local Similarity 99.5%; Pred. No. 5.2e-151;
Matches 852; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 420 CCCTCCCCAGGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACGACGTGGGCGCT 479
Db 3448 CCGCCCGCAGGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACGACGTGGGCGCT 3507
QY 480 GGGCTGGGACTTCATGACGCGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAGAGCT 539
Db 3508 GGGCTGGGACTTCATGACGCGGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAGAGCT 3567
QY 540 GCTGCTGTTTCGACCGCGAGTGCAGCGCGCTGAAGCTGGGCGGCTTCCGCGATGACGCGCG 599
Db 3568 GCTGCTGTTTCGACCGCGAGTGCAGCGCGCTGAAGCTGGGCGGCTTCCGCGATGACGCGCG 3627
QY 600 CGTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 659
Db 3628 CGTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 3687
QY 660 CAGGCGGCGCGCGCGCGAGCGGCTGTGTGACCGGCGGAGCTGTGGGCGCTTCCG 719
Db 3688 CAGGCGGCGCGCGCGCGAGCGGCTGTGTGACCGGCGGAGCTGTGGGCGCTTCCG 3747
QY 720 CGTGCCTCATCTTCGCGTGTCTACCGCGCACTTCCCGTGGGAGCGCGCTGCGGCGCGCG 779
Db 3748 CGTGCCTCATCTTCGCGTGTCTACCGCGCACTTCCCGTGGGAGCGCGCTGCGGCGCGCG 3807
QY 780 CGCCTTCTTCGAGAGTTCGTGTGCGTGAAGCGGCGCGCTTCCCGGGGCTGCTTTCGCA 839
Db 3808 CGCCTTCTTCGAGAGTTCGTGTGCGTGAAGCGGCGCGCTTCCCGGGGCTGCTTTCGCA 3867
QY 840 GTGGCGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGSCCTGAGAGCC 899
Db 3868 GTGGCGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGSCCTGAGAGCC 3927
QY 900 CAGAGCGCGCGCGCGCGCGAGGAGGTGTTCGCTTCCCTCAAGCAGAGCTCACGCTCCGA 959
Db 3928 CAGAGCGCGCGCGCGCGCGAGGAGGTGTTCGCTTCCCTCAAGCAGAGCTCACGCTCCGA 3987
QY 960 GCTGCGCGCGCGCGCTTGCACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 1019
Db 3988 GCTGCGCGCGCGCGCGCTTGCACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 4047
QY 1020 CCGGCGCACTCGGCTCGAGCGCGCTGGGCGCTCAAGCGGAGCGGTGTGCTGACCGAGCGG 1079
Db 4048 CCGGCGCACTCGGCTCGAGCGCGCTGGGCGCTCAAGCGGAGCGGTGTGCTGACCGAGCGG 4107
QY 1080 CAGCGGCTCCCGCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1139
Db 4108 CAGCGGCTCCCGCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 4167
QY 1140 GTTGCGAGTGCCTGCGCGTGTGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1199
Db 4168 GTTGCGAGTGCCTGCGCGTGTGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 4227
QY 1200 CCGCGCGAGCG 1259
Db 4228 CCGCGCGAGCG 4287
QY 1260 CGAGATCTGCGTCTGA 1275
Db 4288 CGAGATCTGCGTCTGA 4303

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 15:10:03 ; Search time 76 Seconds
(without alignments)
7404.783 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atgagctggctgcccaga.....ccatcgagatctgctgtga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/pctus_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	100.0	2598	4	US-09-816-094-1
2	849.6	66.6	7301	4	US-09-816-094-3
3	97.6	7.7	2181	4	US-09-417-197-70
4	97.6	7.7	2184	4	US-09-417-197-70
5	97.6	7.7	2610	2	US-09-212-771-1
6	97.6	7.7	2610	3	US-09-091-058-1
7	93.4	7.3	1599	3	US-09-256-465-1
8	93.4	7.3	1599	4	US-09-167-322-3
9	82.2	6.4	5228	4	US-09-428-711A-15
10	75.4	5.9	1485	4	US-09-801-876B-1
11	74.4	5.8	1008	4	US-09-394-453-3
12	74.4	5.8	2549	3	US-09-467-082-3
13	74.4	5.8	2549	4	US-09-394-455-5
14	74.4	5.8	2608	4	US-09-394-455-35
15	73.8	5.8	2146	4	US-09-842-307-1
16	72.6	5.7	2902	4	US-09-579-664B-4
17	72.4	5.7	1282	2	US-08-878-989-12
18	72.4	5.7	1282	3	US-09-272-796-12
19	72.4	5.7	1282	4	US-09-016-434-953
20	72	5.6	1619	4	US-09-394-455-14
21	69.6	5.5	2746	2	US-09-016-000-12
22	69	5.4	44377	2	US-08-804-227C-7
23	69	5.4	44377	2	US-08-804-198-1
24	67.8	5.3	722	3	US-08-998-416-800
25	67.8	5.3	2673	3	US-09-428-711A-1
26	67.8	5.3	12001	1	US-08-458-568A-11
27	67.6	5.3	4257	2	US-08-690-473-1

28	67.6	5.3	4257	3	US-09-253-821A-1	Sequence 1, Appli
29	67.6	5.3	4257	3	US-08-843-659-1	Sequence 1, Appli
30	67.4	5.3	2908	4	US-09-930-181-1	Sequence 1, Appli
31	66.6	5.2	5207	4	US-09-853-664A-1	Sequence 1, Appli
32	66.2	5.2	3995	4	US-09-423-711A-13	Sequence 13, Appli
33	66.2	5.2	441529	3	US-09-103-840A-1	Sequence 1, Appli
34	64.9	5.1	3228	4	US-09-579-664B-5	Sequence 5, Appli
35	64.6	5.1	1965	4	US-09-252-991A-9230	Sequence 9230, Ap
36	64.6	5.1	2805	4	US-09-252-991A-9064	Sequence 9064, Ap
37	64.6	5.1	3364	4	US-09-930-181-3	Sequence 3, Appli
38	64.4	5.1	2132	2	US-09-159-385-3	Sequence 3, Appli
39	64.4	5.1	2132	3	US-09-186-277-3	Sequence 3, Appli
40	64.4	5.1	4403785	3	US-09-103-840A-2	Sequence 2, Appli
41	64	5.0	1788	4	US-09-417-197-68	Sequence 68, Appli
42	64	5.0	2211	4	US-09-354-455-39	Sequence 39, Appli
43	62.6	4.9	999	4	US-09-252-991A-5206	Sequence 5206, Ap
44	62.6	4.9	1710	4	US-09-252-991A-5271	Sequence 5271, Ap
45	62.4	4.9	3471	2	US-08-715-568A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-816-094-1
; Sequence 1, Application US/09816094
; Patent No. 6534299
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Human
; US-09-816-094-1

Query Match					100.0%; Score 1275; DB 4; Length 2598;
Best Local Similarity					100.0%; Pred. No. 8.3e-231;
Matches 1275; Conservative					0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGAGCGTGGGCTGCCAGAGCCTTGAGCGCCCGCTCCCTGACCTCTGTGGCGCGGG	60		
Db	774	ATGAGCGTGGGCTGCCAGAGCCTTGAGCGCCCGCTCCCTGACCTCTGTGGCGCGGG	833		
Qy	61	ACTGCCCTGGGCTGGTGGCGGTGTCCTTCTCACTGAAGACATGCGAGCCCTGACT	120		
Db	834	ACTGCCCTGGGCTGGTGGCGGTGTCCTTCTCACTGAAGACATGCGAGCCCTGACT	893		
Qy	121	CTCCGCACTGGCGCCGACGAGCTGCAAGCACTACGAACTAGTCCGGAGCTGGG	180		
Db	894	CTCCGCACTGGCGCCGACGAGCTGCAAGCACTACGAACTAGTCCGGAGCTGGG	953		
Qy	181	AAAGGCACTATGGGAGGTGACCTGGTCTACAGGCGACAGCAAAAATGCA	240		
Db	954	AAAGGCACTATGGGAGGTGACCTGGTCTACAGGCGACAGCAAAAATGCA	1013		
Qy	241	CTGAAGTTTGTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGAGGTGAGCATC	300		
Db	1014	CTGAAGTTTGTGAACAGAGCAAAACCAAGCTGAAGAACTTCTACGGAGGTGAGCATC	1073		
Qy	301	ACCAAGCTCTCTCCCGCCCTTCATCATCAAGTCTTTGACGTGTTTGTAGACA	360		
Db	1074	ACCAAGCTCTCTCCCGCCCTTCATCATCAAGTCTTTGACGTGTTTGTAGACA	1133		
Qy	361	GAGGACTGTACCTCTTTCGCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATC	420		

Db 290 CAGAAAGSCACACCCCTGGCTGACTGAGCAGCTCCCGAAACCCCGCAGCTCCCTCGGTGGC 349
Qy 280 TTCCTACGGGAGGTAGCATCACCAACAGCTCTCTCCACACCCCTTCATCATCAAGATC 339
Db 350 TTCCTGTACGAGTCTGTGTGGGCTCTCGCTGGCGCGCACTCAGCCATCGTGACGCGC 409
Qy 340 TTTGACGTGTCCTTTGACAGAGGACTGCTACGCTCTTTGCCAGGAGTAGCAGCTGCT 399
Db 410 TACGGCATTTGCACTCGAGTGGGACACTCTTACAGCTTCTTCGACGAGCCGCTCTGCAC 469
Qy 400 GGGACCTGTTTACATCATCCCTCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGC 459
Db 470 GGGAGCTCATGGCTTCATCCAGCCAAAGTGGGCTCCCGCAGCCCGGCTGCAACGC 529
Qy 460 TGTGTGACAGCTGGGCTGGCTGGCTGACTTATGACAGGGGCGGAGCTGCTGACCGC 519
Db 530 TGGCGCCGCCAGCTGGGCTCGGCTCGGCTGGAGTACATCCACGCCCGCGGCTGTACCGG 589
Qy 520 GACATCAAGCCCGAGAGCTGCTGCTGTTCGACCGGAGTCCCGCGGCTGAAGCTGGC 579
Db 590 GACTGAGCGGAGAGCTGCTGCTGTGAGACCGGCTCCCTGAGGACACGCTGAAGCTGAC 649
Qy 580 GACTTGGCATGACGCGCGGCTGGGCTCGGCTCAAGCGGCTGAGCGGACCATCCCT 639
Db 650 GACTTGGCATGACGCGCGGCTCGGCGGAGCTGCTGCGGCTGGCGGCGGCTGACCGC 709
Qy 640 TACAGGGGCTGAGGTGTC---CAGCGGGGCGCGCGGCTGGGCTGGGCTGGAGCAG 696
Db 710 TACAGCGCGCGGAGCTGTGGCGCGCGCGGCTGCGGCTGGGCTGGGCTGGGCTGGG 769
Qy 697 GGGCTGAGAGCTGTGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
Db 770 GCGCTGAGCGCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
Qy 757 TGGGA---GGCGGCTGGGCGCGGAGCTTCTTTCGAGGAGTTCGTGCTGCGTGGCAGCG 813
Db 830 TGGACCGGCGGCTGGCGGAGCGAGCGGCTTCTACGAGGACTTCTTCTGCTGCTGCTGCT 889
Qy 814 GCGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
Db 890 TCGGGCGAGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949
Qy 874 TTCAGCGCTTACTG 888
Db 950 CTTCGCGGGGCTG 964

RESULT 14

US-10-288-798-40
; Sequence 40, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOLEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Valda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 7477062CB1
US-10-288-798-40

Query Match 18.4%; Score 235; DB 12; Length 1156;
Best Local Similarity 57.7%; Pred. No. 4,7e-46;

Matches 459; Conservative 0; Mismatches 330; Indels 6; Gaps 2;

Qy 100 GAAGACATGAGGCGCTGACTCTCCGACACTGCGCCAGCGAGCTCACCAGCACTAC 159
Db 236 GAGGACATGATGAGCTGAGTGTCTAGACCTGTGCGAGCCGAGGTGGAGCTTCTAC 295
Qy 160 GAACTAGTCCGGAGCTGGCAAGGCACTATGGGAAGTTGACCTGGTGGTCTACAAG 219
Db 296 GAGGAAGTGGTCCCTGGCCAGGGTCTGTATGGCGGCTCTCTTGTGTCACCATCGT 355
Qy 220 GGACAGGCAAAAATGGCACTGAAGTTGTGAACHAGAGCAAAACCAAGCTGAAGAAC 279
Db 356 CAGAAAGGCAACCCCTGGCACTGAAGAGCTCCCGAAACCCGACAGCTCCCTCGTGGC 415
Qy 280 TTCCTACGGAGGTGAGCATCACCAACAGGCTCTCTCCAGCCCTTCATCATCAAGGTC 339
Db 416 TTCCTGACAGTTCTGTGTGGGCTCTCGTGGGCGGCACTCAGCCATCTGTGACGCGC 475
Qy 340 TTTGACGTGCTTTTGAGACAGAGGACTGCTAGTCTTTTCCAGGAGTACGACCTGCT 399
Db 476 TACGGCATTTGGCATCGAGTGGGCACTCTCAGACCTTCTGACGAGGCGCTCTGTCAC 535
Qy 400 GGGGACCTGTTGACATCATCCCTCCCGAGTGGGCTCCCTGAGGACACGCTGGAAGCGC 459
Db 536 GGGGACCTCATGGCTTTCATCCAGCCCAAGTGGGCTCCCGCAGCCCGCGGTGACCGC 595
Qy 460 TGTGTGACGAGCTGGGCTCGGCTGGAGCTTCATGACGGGGGCGAGCTGGTGCACCGC 519
Db 596 TGGCGCCGAGCTGGCTCGGCTCGGCTGGAGTATCATCACGCCCGCGCTGTGTACCGG 655
Qy 520 GACATCAAGCCCGAGAACGCTGCTGCTGCTGCAACGGAGTGGCGCCCGCTAAAGTGGCC 579
Db 656 GACCTGAAGCCGGAGAACGCTGCTGCTGCTGCGACCCCGGCTGCGCGGCTTCAAGTGAOC 715
Qy 580 GACTTGGCATGAGCGCCCGGCTGGGCTGCGGCTGAGCGGTGAGCGGACCATCCCT 639
Db 716 GACTTGGCCACACGAGGCTCGCGGAGCGCTGCTGCGCTGGCGGCGCGCCATCCCC 775
Qy 640 TACACGGGCTGAGGTGTC---CAGCGGGGCGCGCGGCTGGCGGCTGGCGGAGCAG 696
Db 776 TACAGGCCCCGAGCTCTCGGCGCCCCCGCGCTCCCGAGGGCTTCCCATTCAGCCC 835
Qy 697 GGGCTGGAAGCTGGGCTTTCGGGCTGCTCATCTTCTGCGTGTCTACCGGCAACTTCCCG 756
Db 836 GCGCTGGAAGCTGGGCTGGGCTGCTGCTCTTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCT 895

Db	3568	GCTGCTTT	CGACCGAGTGCCGCGCGGTAAGAGCTGCGCCGACTTGGCATGACGGCCG	3627
QY	600	CTGGGGTGC	CGCGTCAAGCGCGTGAGCGGCACCATCCCTTACACGCGCGCTGAGGTGTG	659
Db	3628	CTGGGGTGC	CGCGTCAAGCGCGTGAGCGGCACCATCCCTTACACGCGCGCTGAGGTGTG	3687
QY	660	CGAGGCGGGCGCGCGACGCGGTGGCGGTGGAGACGCGGCGTGGACGTGTGGGCTTCGG	719	
Db	3688	CGAGGCGGGCGCGCGACGCGGTGGCGGTGGAGACGCGGCGTGGAGCGTGTGGGCTTCGG	3747	
QY	720	CGTGCTCATCTTCTGCGTGTCAACCGCAACTTCCCGTGGAGCGCGCTCGGCGCCCGA	779	
Db	3748	CGTGCTCATCTTCTGCGTGTCAACCGCAACTTCCCGTGGAGCGCGCTCGGCGCCCGA	3807	
QY	780	CGCGCTTT	CGAGAGTTCGTGCGTGGCGAGCGGGCCGCTCCGCGGGGTGCTCTTCGCA	839
Db	3808	CGCGCTTT	CGAGAGTTCGTGCGTGGCGAGCGGGCCGCTCCGCGGGGTGCTCTTCGCA	3867
QY	840	GTGGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCACGCGTACTGCGCCCTGAGCC	899	
Db	3868	GTGGCGCGCTTACCGAGCGCGCGCTGCGCATGTTCACGCGTACTGCGCCCTGAGCC	3927	
QY	900	CGAGCGCGCGCGCCAGCCAAGGAGGTTCCTCGCTTCTCAAGCACGAGCTCACGTCCGA	959	
Db	3928	CGAGCGCGCGCGCCAGCCAAGGAGGTTCCTCGCTTCTCAAGCACGAGCTCACGTCCGA	3987	
QY	960	GCTGCGCGCGCGGCGCTTGCACGCGCGCGCAAGCCCGCCGGGACCGCGCGCGCGCCG	1019	
Db	3988	GCTGCGCGCGCGGCGCTTGCACGCGCGCGCAAGCCCGCCGGGACCGCGCGCGCGCCG	4047	
QY	1020	CGGGCCACTCGCGCTCGAGCGCTTGGCGCGCTCAGCGGACGGTGTGTGACGAGCGG	1079	
Db	4048	CGGGCCACTCGCGCTCGAGCGCGCTTGGCGCGCTCAGCGGACGGTGTGTGACGAGCGG	4107	
QY	1080	CAGCGGCTCCGGGCGCGCGCGCGCGCGCGTGGGTGCGTCCCTTGCGCGTGGCGTGC	1139	
Db	4108	CAGCGGCTCCGGGCGCGCGCGCGCGCGCGTGGGTGCGTCCCTTGCGCGTGGCGTGC	4167	
QY	1140	GTTGCCAGTGCCTGCGGTGCTGTGTCGAGACCGCGGCTAGCTCCCGAGGGGCGCCC	1199	
Db	4168	GTTGCCAGTGCCTGCGGTGCTGTGTCGAGACCGCGGCTAGCTCCCGAGGGGCGCCC	4227	
QY	1200	CGGCGGACCGAGCGCGCGGACCAAGAGCAAGAGGACAGTGGTGTGCGCACGCGCAT	1259	
Db	4228	CGGCGGACCGAGCGCGCGGACCAAGAGCAAGAGGACAGTGGTGTGCGCACGCGCAT	4287	
QY	1260	CGAGATCTGCTGTGA	1275	
Db	4288	CGAGATCTGCTGTGA	4303	

```

RESULT 11
US-10-233-613-3
; Sequence 3, Application US/10233613
; Publication No. US20030022339A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536DIV
; CURRENT APPLICATION NUMBER: US/10/233,613
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-233-613-3

```

[illegible]

```

RESULT 12
US-09-861-893-38/c
; Sequence 38, Application US/09861893
; Patent No. US20020045257A1
; GENERAL INFORMATION:
; APPLICANT: Feinberg, Andrew
; APPLICANT: Strichman-Almashanu, Liora
; APPLICANT: Jiang, Shan
; TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND

```

QY	955	TCCAGAGTCGCGCGCGCGCCCTCGACCGCGCGCGCAAGCCCCCGGGGACCCCGCGCCC	1014
Db	721	TCCAGAGTCGCGCGCGCGCCCTCGACCGCGCGCGCAAGCCCCCGGGGACCCCGCGCCC	780
QY	1015	GC CGCGGGGCACCTGCGCCCTCGAGGCGCCTGGGCGCTCAAGCGGACGGTGCTGACCGAG	1074
Db	781	GC CGCGGGGCACCTGCGCCCTCGAGGCGCCTGGGCGCTCAAGCGGACGGTGCTGACCGAG	840
QY	1075	AGCGGCAGCGGCTCCCGGCGCGGCGCCCGCGCGCTCGGTCGGTGCCCTTGCCCGGTGCCG	1134
Db	841	AGCGGCAGCGGCTCCCGGCGCGGCGCCCGCGCGCTCGGTCGGTGCCCTTGCCCGGTGCCG	900
QY	1135	GTGCGGCTGTCAGTGCCCGTGCGCGTGCCCTGTGCCGAGCGCGGCGCTAGCTCCCCAGGGG	1194
Db	901	GTGCGGCTGTCAGTGCCCGTGCGCGTGCCCTGTGCCGAGCGCGGCGCTAGCTCCCCAGGGG	960
QY	1195	CCCCCGCGCGGACCGAGCGCGCGCGGACGAGGCAAGGCGAGGTGGTGCTGCGCCACG	1254
Db	961	CCCCCGCGCGGACCGAGCGCGCGCGGACGAGGCAAGGCGAGGTGGTGCTGCGCCACG	1020
QY	1255	GCCATCGAGATCTCGGCTCTGA	1275
Db	1021	GCCATCGAGATCTCGGCTCTGA	1041

```

RESULT 9
US-09-816-094-3
; Sequence 3, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Human
; US-09-816-094-3

```

[illegible]

```

RESULT 10
US-09-734-032-3
; Sequence 3, Application US/09734032
; Patent No. US20020103116A1
; GENERAL INFORMATION:
; APPLICANT: WEI et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; AND USES
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00536
; CURRENT APPLICATION NUMBER: US/09/734,032
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60205228
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: human
US-09-734-032-3

```

```
Db 284 AGTTTGTGAACAAGAGAAACCAAGCTGAGAACTTCTTACGGAGGTGACATCACCA 343
QY 305 ACAGCCTCTCTCAGCCCTTCTCATCAAGAGTCTTTGAGTGGTCTTTGAGACAGAGG 364
Db 344 ACAGCCTCTCTCAGCCCTTCTCATCAAGAGTCTTTGAGTGGTCTTTGAGACAGAGG 403
QY 365 ACTGCTAGCTTTTGCACAGAGTACGACCTGTGGGGAGCTTTTGACATCATCCCTC 424
Db 404 ACTGCTAGCTTTTGCACAGAGTACGACCTGTGGGGAGCTTTTGACATCATCCCTC 463
QY 425 CCCAGTGGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGCACAGCTGGGGCTGGCGC 484
Db 464 CCCAGTGGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGCACAGCTGGGGCTGGCGC 523
QY 485 TGGACTTATGACAGGGCGGAGCTGTGTGCAACGGGACATCAAGCCCGAGAACGTGTGTC 544
Db 524 TGGACTTATGACAGGGCGGAGCTGTGTGCAACGGGACATCAAGCCCGAGAACGTGTGTC 583
QY 545 TGTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGACTTCGGCATACAGCGCGCGGTGG 604
Db 584 TGTTCGACCGGAGTGGCGCGCTAAAGCTGGCGGACTTCGGCATACAGCGCGCGGTGG 643
QY 605 GCTGCCCGCTCAAGCGGTGAGCGGCACCATCCCTTACACGGCGCCTGAGGTGTGCCAGG 664
Db 644 GCTGCCCGCTCAAGCGGTGAGCGGCACCATCCCTTACACGGCGCCTGAGGTGTGCCAGG 703
QY 665 CGGGCCGCGGACGGGGTGGGTGACACGGCGGTGGAGCTGTGGGCTTTCGGCGTGC 724
Db 704 CGGGCCGCGGACGGGGTGGGTGACACGGCGGTGGAGCTGTGGGCTTTCGGCGTGC 763
QY 725 TCATCTTCTGTGCTCACCAGCAACTTCCTCGTGGGAGGGCGTCCGGCGCGCAGCCCT 784
Db 764 TCATCTTCTGTGCTCACCAGCAACTTCCTCGTGGGAGGGCGTCCGGCGCGCAGCCCT 823
QY 785 TCTTCGAGGAGTTCGTGCGTGGCAGCGGGCGCCCTTCCGGGGCTGCTTCGACGTGGC 844
Db 824 TCTTCGAGGAGTTCGTGCGTGGCAGCGGGCGCCCTTCCGGGGCTGCTTCGACGTGGC 883
QY 845 GCGGCTTACGAGCCCGCTGCGATGTTCCAGCGCTTACTGGCCCTGGAGCCGAGC 904
Db 884 GCGGCTTACGAGCCCGCTGCGATGTTCCAGCGCTTACTGGCCCTGGAGCCGAGC 943
QY 905 GCGCGCGCCAGCAAGAGGTGTTCCGCTTCCTCAAGCAGAGCTCACGTCGAGCTGC 964
Db 944 GCGCGCGCCAGCAAGAGGTGTTCCGCTTCCTCAAGCAGAGCTCACGTCGAGCTGC 1003
QY 965 GCGCGCGCCCTCGCACCGCGCGCAAGCGCCCGGGGACCGCCCGCGCGCGCGGGC 1024
Db 1004 GCGCGCGCCCTCGCACCGCGCGCAAGCGCCCGGGGACCGCCCGCGCGCGCGGGC 1063
QY 1025 CACTGCGCTCGAGGCGCTGGGCGCTCAAGCGGACGGTGTGACCGAGAGCGGACGCG 1084
Db 1064 CACTGCGCTCGAGGCGCTGGGCGCTCAAGCGGACGGTGTGACCGAGAGCGGACGCG 1123
QY 1085 GCTCCGCGCGCGCGCCCGCTCGGTGCGGTGCGCTTCCCGTCCCGGTGCGCGTGC 1144
Db 1124 GCTCCGCGCGCGCGCCCGCTCGGTGCGGTGCGCTTCCCGTCCCGGTGCGCGTGC 1183
QY 1145 CAGTGCCTGTCGGGTGCTGTGCCCGAGCCCGGCTAGTTCCTCCAGGGGCCCCCGGGC 1204
Db 1184 CAGTGCCTGTCGGGTGCTGTGCCCGAGCCCGGCTAGTTCCTCCAGGGGCCCCCGGGC 1243
QY 1205 GGACCGAGCGCGCGGACAGAGCAAGAGCAAGGCAAGGTGTGCTGGCCACCGCCATCAGA 1264
Db 1244 GGACCGAGCGCGCGGACAGAGCAAGAGCAAGGCAAGGTGTGCTGGCCACCGCCATCAGA 1303
QY 1265 TCTGCTCTGA 1275
Db 1304 TCTGCTCTGA 1314
```

RESULT 8

```
US-09-916-790-3
; Sequence 3, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-916-790-3
```

```
Query Match 81.6%; Score 1041; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 2.7e-236;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 ATGGCACTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTG 294
Db 1 ATGGCACTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTG 60
QY 295 AGCATCACCAACAGCCTCTCTCCAGCCCTTTCATCATCAAGCTCTTGTGAGTGGTCTTT 354
Db 61 AGCATCACCAACAGCCTCTCTCCAGCCCTTTCATCATCAAGCTCTTGTGAGTGGTCTTT 120
QY 355 GAGACAGAGACTGCTACAGTCTTTTGGCCAGGAGTACGACCTGCTGGGGACCTGTTTGAC 414
Db 121 GAGACAGAGACTGCTACAGTCTTTTGGCCAGGAGTACGACCTGCTGGGGACCTGTTTGAC 180
QY 415 ATCATCCTTCCAGGTGGGCTCCCTGAGGACACCGTGAAGCGCTGTGTGACGAGCTG 474
Db 181 ATCATCCTTCCAGGTGGGCTCCCTGAGGACACCGTGAAGCGCTGTGTGACGAGCTG 240
QY 475 GGCTGGCGCTGACATTCATGACCGGGCGGCGCTGGTGCACCGGACATCAAGCCCGAG 534
Db 241 GGCTGGCGCTGACATTCATGACCGGGCGGCGCTGGTGCACCGGACATCAAGCCCGAG 300
QY 535 AAGCTGCTGTTCGACACCGGAGTGGCGCGTAAAGCTGGCCGACTTCGCGCATGACG 594
Db 301 AAGCTGCTGTTCGACACCGGAGTGGCGCGTAAAGCTGGCCGACTTCGCGCATGACG 360
QY 595 CGCGCGTGGGCTGCGCGCTCAAGCGCTGAGCGGCAACATCCCTTACACGGCGCTGAG 654
Db 361 CGCGCGTGGGCTGCGCGCTCAAGCGCTGAGCGGCAACATCCCTTACACGGCGCTGAG 420
QY 655 GTCTGACGCGCGCGCGCGCGCTGGCGGTGACACCGCGCTGACGCTGTGGGCC 714
Db 421 GTCTGACGCGCGCGCGCGCGCTGGCGGTGACACCGCGCTGACGCTGTGGGCC 480
QY 715 TTCCGGGTGCTCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGAGGGCGGCTGGGGC 774
Db 481 TTCCGGGTGCTCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGAGGGCGGCTGGGGC 540
QY 775 GCGGACGCTTCTTCGAGGAGTTCGTGGCTGGCAGCGGGCGCTGCGGGGCTGCCT 834
Db 541 GCGGACGCTTCTTCGAGGAGTTCGTGGCTGGCAGCGGGCGCTGCGGGGCTGCCT 600
QY 835 TCCGAGTGGCGCGCTTCACCGGACCGCGCTGGCGGTGACACCGCGCTTACGCGCTG 894
Db 601 TCCGAGTGGCGCGCTTCACCGGACCGCGCTGGCGGTGACACCGCGCTTACGCGCTG 660
QY 895 GAGCCCGAGCGCGCGCGCGCGCAAGAGAGGTGTTCCGCTTCCCTCAAGCAGAGCTCACG 954
Db 661 GAGCCCGAGCGCGCGCGCGCGCAAGAGAGGTGTTCCGCTTCCCTCAAGCAGAGCTCACG 720
```

OTHER INFORMATION: Incyte ID No. US20030207299A1 7312543CB1
US-10-288-798-25

		Query Match	99.98;	Score 1273.4;	DB 12;	Length 2060;
		Best Local Similarity	99.98;	Pred. No. 4.le-291;		
		Matches 1274; Conservative	0;	Mismatches 11;	Indels	0; Gaps
Qy	1	ATGAGCGTGGCGTGCCACAGAGCTGAGCCGCCCGCCTCCCTGCACCTGCTGTGGCGCGGGG	60			
Dd	445	ATGAGCGTGGCGTGCCACAGAGCTGAGCCGCCCGCCTCCCTGCACCTGCTGTGGCGCGGGG	504			
Qy	61	ACTGCCCTTGGGCTGTGGCGTGTGGCGTTCTCTAATAAGACATGCAAGGCCCTGAATT	120			
Dd	505	ACTGCCCTTGGGCTGTGGCGTGTGGCGTTCTCTAATAAGACATGCAAGGCCCTGAATT	564			
Qy	121	CTCCGACACTGGCGCGCACGAGCGTACCAGAAGCATACGAATAGTTCGGGAGCTGGGC	180			
Dd	565	CTCCGACACTGGCGCGCACGAGCGTACCAGAAGCATACGAATAGTTCGGGAGCTGGGC	624			
Qy	181	AAGGCACCTATGGGAAGTTTGACCTGGTGGTCTACAAGGGCAAGGCAAAATAATGGCA	240			
Dd	625	AAGGCACCTATGGGAAGTTTGACCTGGTGGTCTACAAGGGCAAGGCAAAATAATGGCA	684			
Qy	241	CTGAAGTTTGTGAACAAGAGCAAACCAAGCTGAAGAACCTTCCTACGGGAGGTGAGCATC	300			
Dd	685	CTGAAGTTTGTGAACAAGAGCAAACCAAGCTGAAGAACCTTCCTACGGGAGGTGAGCATC	744			
Qy	301	ACCAAGCCTCTCCTCCAGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	360			
Dd	745	ACCAAGCCTCTCCTCCAGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	804			
Qy	361	GAGGACTGCTACGCTTTTGGCCAGGAGTAGCACCTGCTGGGACCTGTTTGACATCATC	420			
Dd	805	GAGGACTGCTACGCTTTTGGCCAGGAGTAGCACCTGCTGGGACCTGTTTGACATCATC	864			
Qy	421	CCTCCCAGGTGGGCGTCCCTTGAGACAACGGTGAAGCGCTGTGTGACGAGCTGGGCGTG	480			
Dd	865	CCTCCCAGGTGGGCGTCCCTTGAGAGACAACGGTGAAGCGCTGTGTGACGAGCTGGGCGTG	924			
Qy	481	GCGCTGACTTCATGACGGGCGGACGCTGTGTGACCGGACATCAAGCCGAGAACGTG	540			
Dd	925	GCGCTGACTTCATGACGGGCGGACGCTGTGTGACCGGACATCAAGCCGAGAACGTG	984			
Qy	541	CTGCTGTTTCGACCGCGAGTGC CGCGCGTAAAGCTGGCCGACTTCGGCGATGACGCGCGC	600			
Dd	985	CTGCTGTTTCGACCGCGAGTGC CGCGCGTAAAGCTGGCCGACTTCGGCGATGACGCGCGC	1044			
Qy	601	GTGGGCTGCCGCTCAAGCGGCTGAGCGGACCATCCCTTACACGGCGCCTGAGGTGTC	660			
Dd	1045	GTGGGCTGCCGCTCAAGCGGCTGAGCGGACCATCCCTTACACGGCGCCTGAGGTGTC	1104			
Qy	661	CAGCGGCGCGCGCACGGGCTGGCGTGGACACGGGCGTGGACGCTGTGGGCCCTTCGCG	720			
Dd	1105	CAGCGGCGCGCGCACGGGCTGGCGTGGACACGGGCGTGGACGCTGTGGGCCCTTCGCG	1164			
Qy	721	GTGCTCATCTTCGCTGCTCACCGGCACTTCGGTGGGAGGCGGCGTGGGCGCGCGAC	780			
Dd	1165	GTGCTCATCTTCGCTGCTCACCGGCACTTCGGTGGGAGGCGGCGTGGGCGCGCGAC	1224			
Qy	781	GCTTCTTCAGGAGTTCTGCTGGTGGCAGCGGGCGCCCTTCCCGGGCTGCTTCGCG	840			
Dd	1225	GCTTCTTCAGGAGTTCTGCTGGTGGCAGCGGGCGCCCTTCCCGGGCTGCTTCGCG	1284			
Qy	841	TGGCGCGCTTTCACGAGCCCGCGCTGCGATGTTTCAGCGCTTACTTGGGCCCTGGAGGCC	900			
Dd	1285	TGGCGCGCTTTCACGAGCCCGCGCTGCGATGTTTCAGCGCTTACTTGGGCCCTGGAGGCC	1344			
Qy	901	GAGCGCGCGGCCACGACGAGGAGTGTTCGCTTCTCAAGCACGAGCTCACGTCGCG	960			
Dd	1345	GAGCGCGCGGCCACGACGAGGAGTGTTCGCTTCTCAAGCACGAGCTCACGTCGCG	1404			
Qy	961	CTGCGCGCGGCCCTTCGACCGCGCGCAAGCCCCCGGGGACCGCGCGCGCGCC	1020			

```

RESULT 7
US-09-916-790-1
; Sequence 1, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374. NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (274)...(1314)
; NAME/KEY: misc.feature
; LOCATION: (1)...(2893)
; OTHER INFORMATION: n = A,T,C or G
US-09-916-790-1

```

Query Match	88.8%	Score 1131.8	DB 9	Length 2893	
Best Local Similarity	94.8%	Pred. NC. 1.1e-257			
Matches 1205	Conservative	0	Mismatches 57	Indels 9	Gaps 3
Qy	14	GCCACAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCGCGGGGACTGCCCTCGGGC	73		
Db	44	GCCTTGAACCGGCCCGGCTCCTTGACCTTGCTTGTGGCGCGAAGACTGCCCTCGGGC	103		
Qy	74	CTGGTCCGCG--GTGTGCCCTTCTCATCTGAAGACATGCAAGGCCCTGACTCTCTCGCACAC	130		
Db	104	CTGGGTCCGGGTGTGCCCTTCTCATCTTGAAGACATGCAAGGCCCTGACTCTCTCGCACAC	163		
Qy	131	TGG--CGCGCAGCGAGCTTCAACCAAGCACTACGAACTAGT-----CGGGGAGCTGGGCAAG	184		
Db	164	TTGCCCGCCAGCGAGCTCACCCACGACCTACGAACACTAGTCTCGGGAGCTGGGCAAG	223		
Qy	185	GCACCTATGGGAAGTTGACCTGGTGTCTACAAGGGCAAGGCAACAAAATGGCACTGA	244		
Db	224	GCACCTACGGGAAGTTGATCTGGTGTCTACAAGGGCAAGGCAACAAAATGGCACTGA	283		
Qy	245	AGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTTACGGGAGGTGAGCATCACCA	304		

```
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: human
US-09-734-032-1

Query Match      99.9%; Score 1273.4; DB 10; Length 1275;
Best Local Similarity 99.9%; Pred. No. 3.9e-291;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTCTGCTGGCGCGGGG 60
Db 1 ATGAGCTGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTCTGCTGGCGCGGGG 60

QY 61 ACTGCCCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 120
Db 61 ACTGCCCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 120

QY 121 CTCGGACACTGGCGCGCAGCGAGCTCACCAGCACTACGAATAGTCCCGGAGCTGGGC 180
Db 121 CTCGGACACTGGCGCGCAGCGAGCTCACCAGCACTACGAATAGTCCCGGAGCTGGGC 180

QY 181 AAAGGCACCTATGGGAAGGTTGACCTGGTGGTCTTACAGGSCACAGSCACAAAATGGCA 240
Db 181 AAAGGCACCTATGGGAAGGTTGACCTGGTGGTCTTACAGGSCACAGSCACAAAATGGCA 240

QY 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC 300
Db 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC 300

QY 301 ACCAACAGCTTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA 360
Db 301 ACCAACAGCTTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA 360

QY 361 GAGGACTGTACTGTTTGGCCAGAGTACGACACTGCTGGGACCTGTTTGACATCATC 420
Db 361 GAGGACTGTACTGTTTGGCCAGAGTACGACACTGCTGGGACCTGTTTGACATCATC 420

QY 421 CCTTCCAGGTGGGCTCCTTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCTG 480
Db 421 CCTTCCAGGTGGGCTCCTTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCTG 480

QY 481 GCGTGGAATTCATGACGCGCGGAGCTGTGTGACCGGACATCAAGCCGCGAGAGCTG 540
Db 481 GCGTGGAATTCATGACGCGCGGAGCTGTGTGACCGGACATCAAGCCGCGAGAGCTG 540

QY 541 GTGCTGTTCACCGCGAGTGGCGCGGTAAGCTGGCGGACTTCGCGATGACGCGCGCG 600
Db 541 GTGCTGTTCACCGCGAGTGGCGCGGTAAGCTGGCGGACTTCGCGATGACGCGCGCG 600

QY 601 GTGGGCTGCCGCGTCAAGCGCGTGAAGCGGACCATCTTACAGCGCGCTGAGGTGTC 660
Db 601 GTGGGCTGCCGCGTCAAGCGCGTGAAGCGGACCATCTTACAGCGCGCTGAGGTGTC 660

QY 661 CAGCGCGCGCGCGCGAGCTGGCGGTGGACACAGCGGCGTGGACGTGTGGGCTTCGCG 720
Db 661 CAGCGCGCGCGCGCGAGCTGGCGGTGGACACAGCGGCGTGGACGTGTGGGCTTCGCG 720

QY 721 GTGCTCATCTTCTGCTGCTCACCGGCACTTCCCGTGGAGGCGGCTGGGCGCGGAC 780
Db 721 GTGCTCATCTTCTGCTGCTCACCGGCACTTCCCGTGGAGGCGGCTGGGCGCGGAC 780

QY 781 GCCTTCTTCAGAGATTGCTGGCTGGCAGCGGCGCTGCGGCGCTGCTTGGCAG 840
Db 781 GCCTTCTTCAGAGATTGCTGGCTGGCAGCGGCGCTGCGGCGCTGCTTGGCAG 840

QY 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCAGCGCTTACTTGGCCCTGGAGCCC 900
Db 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCAGCGCTTACTTGGCCCTGGAGCCC 900

QY 901 GAGGCGCGCGCGCGCGAGGAGGTTTCCGCTTCTCAAGCAGAGCTCAAGTCCGAG 960
Db 901 GAGGCGCGCGCGCGCGAGGAGGTTTCCGCTTCTCAAGCAGAGCTCAAGTCCGAG 960

RESULT 6
US-10-288-798-25
; Sequence 25, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B.;
; APPLICANT: WALIA, Narinder K.; HAFADIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: RAMKUMAR, Jayalaxmi;
; APPLICANT: KEARNEY, Liam; POLICKY, Warren, Bridget A.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil;
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 2060
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```


QY	601	GTGGGCTGCCCGCTCAAGCGCGT	GAGCGGCA	CCATCCCTTACACGGCGCCTGAGGTGCG	660
Db	717	GTGGGCTGCCCGCTCAAGCGCGT	GAGCGGCA	CCATCCCTTACACGGCGCCTGAGGTGCG	776
QY	661	CAGCGGGCGCGCGCGCGGCT	TGGCGTGACACGGCGGTGGACGTG	TGGGCGCTTCGGC	720
Db	777	CAGCGGGCGCGCGCGGCT	TGGCGTGACACGGCGGTGGACGTG	TGGGCGCTTCGGC	836
QY	721	GTGCTCATCTTCTCGTGCT	CACCGGCAATCCGTGGGAGCGCGT	TGGGCGCGCGAC	780
Db	837	GTGCTCATCTTCTCGTGCT	CACCGGCAATCCGTGGGAGCGCGT	TGGGCGCGCGAC	896
QY	781	GCCCTCTTCAGGAGTTCTGCGCT	TGCAGCGGGCGCCCTGCGGGCTGCTTCGAG	840	
Db	897	GCCCTCTTCAGGAGTTCTGCGCT	TGCAGCGGGCGCCCTGCGGGCTGCTTCGAG	956	
QY	841	TGGCGCGCTTACCGAGCCCGCT	CGCATGTTCCAGCGCTTACTTGGGCCCTGGAGCCC	900	
Db	957	TGGCGCGCTTACCGAGCCCGCT	CGCATGTTCCAGCGCTTACTTGGGCCCTGGAGCCC	1016	
QY	901	GAGCGCGCGCGCCAGCCAAAGGAGT	TTCGGCTTCTCAAGCACGAGCTCA	CGTCCGAG	960
Db	1017	GAGCGCGCGCGCCAGCCAAAGGAGT	TTCGGCTTCTCAAGCACGAGCTCA	CGTCCGAG	1076
QY	961	CTGCGCGCGCGGCCCTCGCAC	CCGCGCGCAAGACCCCGGGGACCCCGCGCGCGCC	1020	
Db	1077	CTGCGCGCGCGGCCCTCGCAC	CCGCGCGCAAGACCCCGGGGACCCCGCGCGCGCC	1136	
QY	1021	GGGCACTGCGCTCGAGGGCGCT	TGGGCCCGTCAAGCGGACGGTGTG	ACCGAGAGCGC	1080
Db	1137	GGGCACTGCGCTCGAGGGCGCT	TGGGCCCGTCAAGCGGACGGTGTG	ACCGAGAGCGC	1196
QY	1081	AGGCGCTCCCGGGCGCGCGCCCGCGCT	CGGGTCTGCTTCCCGGCGCGGTGCGC	CGCGGTGCGC	1140
Db	1197	AGGCGCTCCCGGGCGCGCGCCCGCGCT	CGGGTCTGCTTCCCGGCGCGGTGCGC	CGCGGTGCGC	1256
QY	1141	GTGCCAGTGCCGCTGCCGCTG	TGCCCGAGCCCGGCTAGCTCCCGAGGGCGCCCC	1200	
Db	1257	GTGCCAGTGCCGCTGCCGCTG	TGCCCGAGCCCGGCTAGCTCCCGAGGGCGCCCC	1316	
QY	1201	GGCCGACACAGCGCGCGGCA	CAAGAGCAAGGCGAGTGTGTGGCCACGGGCATC	1260	
Db	1317	GGCCGACACAGCGCGCGGCA	CAAGAGCAAGGCGAGTGTGTGGCCACGGGCATC	1376	
QY	1261	GAGATCTGCGTCTGA		1275	
Db	1377	GAGATCTGCGTCTGA		1391	

```

RESULT 3
US-09-816-094-1
; Sequence 1, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Human
US-09-816-094-1

```

Query Match	100.0%	Score 1275;	DB 9;	Length 2598;
Best Local Similarity	100.0%;	Pred. No. 1.7e-291;		
Matches 1275;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGAGCGTGGCGTCCGACAGACCTGAGCGCGCCCGCTCCCTGAACTGCTGTGGGCCGCGG	60
Db	774	ATGAGCGTGGCGTCCGACAGACCTGAGCGCGCCCGCTCCCTGAACTGCTGTGGGCCGCGG	833
QY	61	ACTGCCCTGGCGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGCAGGCCCTGACT	120
Db	834	ACTGCCCTGGCGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGCAGGCCCTGACT	893
QY	121	CTCCGCACACTGGCGCGCAGCGAGCTCAACGAACACTACGAACCTAGTCCGGGAGCTGGGC	180
Db	894	CTCCGCACACTGGCGCGCAGCGAGCTCAACGAACACTACGAACCTAGTCCGGGAGCTGGGC	953
QY	181	AAAGCACCTATGGGAAGTTGACCTGTGGTCTACAAGGCACAGGCACAAAAATGCA	240
Db	954	AAAGCACCTATGGGAAGTTGACCTGTGGTCTACAAGGCACAGGCACAAAAATGCA	1013
QY	241	CTGAAGTTTGAAACAAGACAAACACGCTGAAGACTTCTTACGGGAGGTGAGCATC	300
Db	1014	CTGAAGTTTGAAACAAGACAAACACGCTGAAGACTTCTTACGGGAGGTGAGCATC	1073
QY	301	ACCAACAGCCTCTCCTCCAGGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	360
Db	1074	ACCAACAGCCTCTCCTCCAGGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	1133
QY	361	GAGACTGCTACGCTTTTGGCCAGGAGTACGACCTCTGCTGGGACCTTTTGAACATCATC	420
Db	1134	GAGACTGCTACGCTTTTGGCCAGGAGTACGACCTCTGCTGGGACCTTTTGAACATCATC	1193
QY	421	CCTCCCGAGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACGACAGCTGGGCGTG	480
Db	1194	CCTCCCGAGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACGACAGCTGGGCGTG	1253
QY	481	GCGCTGACCTTCATGACACGGCGGCGACCTGGTGCACCGGCACATCAAGCCCGAGAACGTG	540
Db	1254	GCGCTGACCTTCATGACACGGCGGCGACCTGGTGCACCGGCACATCAAGCCCGAGAACGTG	1313
QY	541	CTGCTGTTGCAACCGCGAGTGC CGCGCGGTAAAGCTGGCCGACTTCGCATGACCGCGCGC	600
Db	1314	CTGCTGTTGCAACCGCGAGTGC CGCGCGGTAAAGCTGGCCGACTTCGCATGACCGCGCGC	1373
QY	601	GTGGGCTGCCGCTCAAGCGCGGTGAGCGGCACATCCCTTACACGGCGCCTGAGGTGTC	660
Db	1374	GTGGGCTGCCGCGTCAAGCGCGGTGAGCGGCACATCCCTTACACGGCGCCTGAGGTGTC	1433
QY	661	CAGCGGCGCGCGCGACGGGCTGGCGGTGGACAACGGGCGTGAACGTGTGGGCTTCGCG	720
Db	1434	CAGCGGCGCGCGCGACGGGCTGGCGGTGGACAACGGGCGTGAACGTGTGGGCTTCGCG	1493
QY	721	GTGCTCATCTTCTCGTGTCTACCGGCACTTCCGTGGGAGCGCGTCCGGCGCGCAC	780
Db	1494	GTGCTCATCTTCTCGTGTCTACCGGCACTTCCGTGGGAGCGCGTCCGGCGCGCAC	1553
QY	781	GCCTTCTTCAGAGTTCGTGCGCTGGCAGCGGGCGCGCTGCGGGGCTGCCCTTGCAG	840
Db	1554	GCCTTCTTCAGAGTTCGTGCGCTGGCAGCGGGCGCGCTGCGGGGCTGCCCTTGCAG	1613
QY	841	TGGCGCGCTTACCGAGCCCGCGTGGCATCTCCAGCGCTTACAGCGCTTACGCCCTTGAAGCCC	900
Db	1614	TGGCGCGCTTACCGAGCCCGCGTGGCATCTCCAGCGCTTACAGCGCTTACGCCCTTGAAGCCC	1673
QY	901	GAGCGCGCGCGCCAGCAAGAGGTGTTCGCTTCTCAAGCACAGCTCACGTCGAG	960
Db	1674	GAGCGCGCGCGCCAGCAAGAGGTGTTCGCTTCTCAAGCACAGCTCACGTCGAG	1733
QY	961	CTGGCGCGCGGCCCTTGACACCGCGGCGGCAAGCCCCCGGGGACCGCCCGCCCGCGC	1020
Db	1734	CTGGCGCGCGGCCCTTGACACCGCGGCGGCAAGCCCCCGGGGACCGCCCGCCCGCGC	1793
QY	1021	GGGCACTGGCTTCAGAGCGCTGGCGGCTCAAGCGACCGGTGTGTACCGAGAGCGC	1080
Db	1794	GGGCACTGGCTTCAGAGCGCTGGCGGCTCAAGCGACCGGTGTGTACCGAGAGCGC	1853
QY	1081	AGGGGCTCCCGCGCGCGCCCCCGCGCTGGGTGGCTTGCCTTGCCTGTGCGGTGCG	1140

Qy	1261	GAGATCTGGGTCTGA	1275
Db	1261	GAGATCTGGGTCTGA	1275
RESULT 2			
US-10-016-985-3			
; Sequence 3, Application US/10016985			
; Publication No. US20020123621A1			
; GENERAL INFORMATION:			
; APPLICANT: Walke, D. Wade			
; APPLICANT: Maricar, Miranda			
; APPLICANT: Yu, Xuanchuan (Sean)			
; APPLICANT: Friddle, Carl Johan			
; TITLE OF INVENTION: No. US20020123621A1e1 Human Kinase and Polynucleotides			
; TITLE OF INVENTION: Encoding the Same			
; FILE REFERENCE: LEX-0273-USA			
; CURRENT APPLICATION NUMBER: US/10/016,985			
; CURRENT FILING DATE: 2001-12-07			
; PRIOR APPLICATION NUMBER: US 60/251,941			
; PRIOR FILING DATE: 2000-12-07			
; NUMBER OF SEQ ID NOS: 3			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 1473			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-016-985-3			
Query Match 100.0%; Score 1275; DB 13; Length 1473;			
Best Local Similarity 100.0%; Pred. No. 1.6e-291;			
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGAGCGTGGGTCCGACAGCGTGAGCCGCCGCGTCCTGTACCTGACTGCTGTGGCCCGGGG	60
Db	117	ATGAGCGTGGGTGCCCCAGAGCCTGAGCCGCCCGCGTCCCTGTGACTGCAAGACATGCAGGCCCTGACT	176
Qy	61	ACTGCCCTGGGCGTGTGCGGTTGCCCTTCTCACTAAAGACATGCAGGCCCTGACT	120
Db	177	ACTGCCCTGGGCGTGTGCGGTTGCCCTTCTCACTAAAGACATGCAGGCCCTGACT	236
Qy	121	CTCGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACTAGTCCGGGAGCTGGGC	180
Db	237	CTCGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACTAGTCCGGGAGCTGGGC	296
Qy	181	AAGGCACCTATTGGGAAGTTTGAACCTGGTGGTCTAAAGGGCAGAGCAAAAATGGCA	240
Db	297	AAAGGCACCTATTGGGAAGTTTGAACCTGGTGGTCTAAAGGGCAGAGCAAAAATGGCA	356
Qy	241	CTGAAGTTTCTGAACAAGAGCAAACCACAGCTGAAGAATCTCTACGGGAGGTGAGCATC	300
Db	357	CTGAAGTTTCTGAACAAGAGCAAACCACAGCTGAAGAATCTCTACGGGAGGTGAGCATC	416
Qy	301	ACCAACAGCCTCTCTCCAGCCCTTCATCATCAAGTCTTTGACGTGGTCTTTGAGACA	360
Db	417	ACCAACAGCCTCTCTCCAGCCCTTCATCATCAAGTCTTTGACGTGGTCTTTGAGACA	476
Qy	361	GAGACTGTACGTCTTTTGCCAGGAGTAGCACCTCTCTGGGACCTGTTTGACATCATC	420
Db	477	GAGACTGTACGTCTTTTGCCAGGAGTAGCACCTCTCTGGGACCTGTTTGACATCATC	536
Qy	421	CCTCCCAGGTGGGCTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCGTG	480
Db	537	CCTCCCAGGTGGGCTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCGTG	596
Qy	481	GCCTGTGACTTCATGCA CGGGCGGACGCTGGTGCACCGCGACATCAAGCCCGAGAACGTG	540
Db	597	GCCTGTGACTTCATGCA CGGGCGGACGCTGGTGCACCGCGACATCAAGCCCGAGAACGTG	656
Qy	541	CTGCTGTTCGACCGCGAGTGC GCCCGCGGTAAAGTGGCCGACATTCAGCCCGCGCG	600
Db	657	CTGCTGTTCGACCGCGAGTGC GCCCGCGGTAAAGTGGCCGACATTCAGCCCGCGCG	716

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 16:12:04 / Search time 334 Seconds
(without alignments)
12687.392 Million cell updates/sec

Title: US-10-016-985-1
Perfect score: 1275
Sequence: 1 atgagctgggtcccgaga.....ccatcgagatctggtctga 1275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/ECT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/ECTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	100.0	1275	13	US-10-016-985-1
2	1275	100.0	1473	13	US-10-016-985-3
3	1275	100.0	2598	9	US-09-816-094-1
4	1275	100.0	2598	14	US-10-233-613-1
5	1273.4	99.9	1275	10	US-09-734-032-1
6	1273.4	99.9	2060	12	US-10-288-798-25
7	1131.8	88.8	2893	9	US-09-916-790-1
8	1041	81.6	1041	9	US-09-916-790-3
9	849.6	66.6	7301	9	US-09-816-094-3
10	849.6	66.6	7301	10	US-09-734-032-3
11	849.6	66.6	7301	14	US-10-233-613-3
12	767.6	60.2	788	9	US-09-861-893-38
13	236.6	18.6	1247	12	US-10-353-690-49
14	235	18.4	1156	12	US-10-288-798-40
15	99.8	7.8	1257	9	US-09-799-875-15
16	99.8	7.8	1260	12	US-10-303-664A-5

Sequence 13, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 43, Appli
Sequence 984, Ap
Sequence 166, App
Sequence 70, Appli
Sequence 138, App
Sequence 3, Appli
Sequence 254, App
Sequence 5, Appli
Sequence 1, Appli
Sequence 2340, Ap
Sequence 6069, Ap
Sequence 1, Appli
Sequence 25154, A
Sequence 1, Appli
Sequence 237, App
Sequence 1416, Ap
Sequence 7945, Ap
Sequence 8220, Ap
Sequence 8319, Ap
Sequence 1416, Ap
Sequence 4992, Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 557, App

99.8 7.8 1826 9 US-09-799-875-13
99.8 7.8 1826 12 US-10-303-664A-4
99.8 7.8 1864 11 US-09-819-607-1
99.2 7.8 2978 11 US-09-955-999-43
98.4 7.7 1854 14 US-10-198-846-9984
98.2 7.7 1972 13 US-10-098-841-166
97.6 7.7 2181 14 US-10-072-036-70
97.6 7.7 2184 14 US-10-072-036-138
97.6 7.7 2610 10 US-09-970-000-3
93.4 7.3 1599 12 US-10-324-985A-3
93.4 7.3 1715 12 US-10-007-926A-254
89.4 7.0 1140 12 US-10-324-985A-5
89.4 7.0 2626 12 US-10-324-985A-1
88.8 7.0 1635 10 US-09-880-107-2340
87.2 6.8 1944 14 US-10-156-761-6069
87.2 6.8 9025608 14 US-10-156-761-1
86.2 6.8 819 14 US-10-275-546-1
86.2 6.8 858 12 US-10-029-386-25154
86.2 6.8 1149 12 US-09-863-776-1
86.2 6.8 1376 10 US-09-764-868-237
86.2 6.8 32192 10 US-09-764-847-1416
86.2 6.8 32192 11 US-09-764-891-7945
86.2 6.8 32192 11 US-09-764-891-8220
86.2 6.8 32192 11 US-09-764-891-8319
86.2 6.8 32192 14 US-10-092-154-1416
85.8 6.7 1155 14 US-10-156-761-4992
85.8 6.7 9025608 14 US-10-156-761-1
85.4 6.7 3270 10 US-09-835-081-1
85.2 6.7 1368 10 US-09-764-868-557

ALIGNMENTS

RESULT 1

US-10-016-985-1
; Sequence 1, Application US/10016985
; Publication No. US2002012362A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Maricar, Miranda
; APPLICANT: Yu, Xuanchuan (Sean)
; APPLICANT: Fridele, Carl Johan
; TITLE OF INVENTION: No. US2002012362A1: Human Kinase and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0273-USA
; CURRENT APPLICATION NUMBER: US/10/016,985
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/251,941
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-985-1

Query Match 100.0%; Score 1275; DB 13; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.6e-291;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCTGGGTGCTCCAGAGCCTGAGCGCGCGCTCTGAGCTCTGTGGGCGCGGG 60
Db 1 ATGAGCTGGGTGCTCCAGAGCCTGAGCGCGCGCTCTGAGCTCTGTGGGCGCGGG 60

Qy 61 ACTCCCTGGCTGTGCGGTGTCCTTCTACTGAAGACATGCGAGCCCTGACT 120
Db 61 ACTCCCTGGCTGTGCGGTGTCCTTCTACTGAAGACATGCGAGCCCTGACT 120

Qy 121 CTCGACACTGGCGCGCGAGCGAGTACCAAGCACTACGAAGTACGAGCTGGG 180
Db 121 CTCGACACTGGCGCGCGAGCGAGTACCAAGCACTACGAAGTACGAGCTGGG 180

gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 135 a 206 c 260 g 165 t 1 others
ORIGIN

Query Match 46.6%; Score 594.4; DB 13; Length 767;
Best Local Similarity 88.8%; Pred. No. 3.3e-105;
Matches 677; Conservative 0; Mismatches 81; Indels 4; Gaps 3;

QY 124 CGCACACTGGCGCGCCAGCGCTCACCAGCACTACGAACTAGTCCGGGAGCTGGGCAAA 183
DB 1 CGCACACTGGCTGCCAGCGAGTGTACCAAGCACTATGAGCTTGTGCGGAGCTGGGCAAA 60

QY 184 GGCACCTATGGGAAGTTGACCTGTGTCTCAAGGGCCACAGGCACAAAATGGCACTG 243
DB 61 GGGACCTACGGGAAGTTCGACCTGTGTCTTACAGGGCCACAGGCACCTAAAATGGCCCTG 120

QY 244 AAGTTGTGAACAAGACAAACCAAGCTGAGAACCTTCTTACGGGAGGTGAGCATCACC 303
DB 121 AAATTTGTGAATAGAGTAAGACAAAGCTGAGAACTTTCTGCGGTGAGGTGAGCATCACC 180

QY 304 AACAGCCCTCTCTCCAGGCCCTTTCATCATCAAGTCTTTGACGTGTCTTTGAGACAGAG 363
DB 181 AACAGCCCTGTGTCTAGCCCTTTCATCATCAAGTCTTTGACGTGTCTTTGAGACCCAG 240

QY 364 GACTGCTACGCTTTTCCAGAGTACGCACTGCTGGGACCTGTTTGACATCATCCCT 423
DB 241 GAGTGTATGTCTTTTGGCCAGAGTATGACCTGCTGGGACCTGTTTGACATCATCCCT 300

QY 424 CCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGCAGACGCTGGGCTGGCG 483
DB 301 CCTCAGGTGGGGCTCCAGAGGACACGGTGAAGCGCTGTGTGCAGACGCTGGGCTGGCG 360

QY 484 CTGGACTTCATGCACGGCGGCGCTGTGTGCACCGCACATCAAGCCCGAGAACGTGCTG 543
DB 361 CTGGACTTCATGCACAGCAGCAGCTGTGTGCACCGCACATCAAGCCCGAGAACGTGCTG 420

QY 544 CTGTTGACCCGCGAGTGCCTGGCTAAAGTGGCCGACTTGGGATGACCGCGCGCTG 603
DB 421 CTGTTGACCCGCGAGTGCCTGGCTGTGAAGCTGGCTGACTTCGGCATGACCGCGCGCTG 480

QY 604 GGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACAGCGGCTGAGGTGTGCCAG 663
DB 481 GGCTGCGCGTGAAGCGGTGAGCGGCACTATACCGTACACAGCACCCGAGGTGTGCCAG 540

QY 664 GCGGGCCGCGCGAGCGGCTGGCGGTGACACGGGCGTGGACGTGTGGGCCCTTCGGCGTG 723
DB 541 GCTGGCGCGCGGATGGCTTCGGGTGACACGGGCGTGGACGTGTGGGCCCTTTGGCGTG 600

QY 724 CTCATCTTCTCGGTGCTCACCGGCACTTCC-CGTGGAGCGGCGTGGCGCGCGAGCGC 782
DB 601 CTCATCTTCTGTGTGCTCACTGGCAACTTCNCGTGGGAGGCTGATCGGCTGGGATGC 660

QY 783 CTTCTTCAGGAGTTCTGTCGCTGACGCGGGCGCTGCGGCGCTGCGGCGCTTCCAGCGCTT 842
DB 661 CTTCTTCAGGAGTTCTGTCGCTGACGCGGGTGCCTGCGGCGCTGCGGCGCTTCCAGCGCTT 719

QY 843 GCGCGCTTTCACCGAGCGCGGCTGCGGCTGCTCCAGCGCTT 884
DB 720 GCGC--GTTTACCGAGCTGCGCTGCGTATGTTACGCGCTT 759

Search completed: December 7, 2003, 16:46:04
Job time : 2030 secs

```

VERSION BQ179888.1 GI:20355380
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .767
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_EW0"
/note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GTCGCTGGGA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 109 a 203 c 242 g 152 t
ORIGIN

Query Match 46.8%; Score 596.8; DB 13; Length 706;
Best Local Similarity 90.5%; Pred. No. 1.1e-105;
Matches 637; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 265 ACCAAGCTGAAGACTTCTCAGGGAGGTGAGCATCACCAACAGCCTCTCCAGCCCC 324
Db 2 ACAAGCTGAAGACTTCTCGGTGAGGTGAGCATCACCAACAGCCTCTCCTAGCCCC 61

QY 325 TTCATCATCAAGGCTTTTGAGCTGGTCTTTGAGACAGAGGACTGCTAGCTTTTGCCCG 384
Db 62 TTCATCATCAAGGCTTTTGAGCTGGTCTTTGAGACAGGAGTGTATGCTTTTGCCCG 121

QY 385 GAGTAGCCACCTGCTGGGGACCTGTTTGACATCATCTCCCTCCAGGTGGGCTCCCTGAG 444
Db 122 GAGTAGCCACCTGCTGGGGACCTGTTTGACATCATCTCCCTCCAGGTGGGCTCCCG 181

QY 445 GACACGGTGAACCGCTGTGCACAGCTGGGCTGGGCTGGGCTGACATCATCAGCGCGG 504
Db 182 GACACGGTGAACCGCTGTGTGACAGAGCTGGGGCTGGGCTGGGCTGACATCATCAGCAGG 241

QY 505 CAGCTGTGTGACCGCGACATCAAGCCCCGAGACCTGTGCTGTGTCGACCGGAGTGGCGG 564

```

```

Db 242 CAGCTGTGTGACCGTGACATCAAGCCCGAGATGTGCTGTGTTGACCGTGTGCGC 301
QY 565 CGGTTAAAGCTGGCCGACTTTCGGCATGACCGCGCGCTGGGCTGCCCGCTCAAGCGGTG 624
Db 302 CGTGTGAAGCTGGCTGACTTTCGGCATGACCGCGCGCTGGGCTGCCCGTGAAGCGGTG 361
QY 625 AGCGGCACCATTCCTTACAGCGGCCCTTACAGGTGTGCCAGGGGCCCGCGCGAGCGGTG 684
Db 362 AGCGGCACCATTCCTTACAGCGGCCCTTACAGGTGTGCCAGGGGCCCGCGCGAGGTG 421
QY 685 GCGTGTGACACGGCGGTGGACGTGTGGCCCTTGGCGGTGCTCATCTCTCGGTGCTCAC 744
Db 422 GCGTGTGACACGGCGGTGGACGTGTGGCCCTTGGCGGTGCTCATCTCTCGGTGCTCAC 481
QY 745 GGCAACTTCCGCTGGGAGGGCGGTGCGCGCGCGCGACGCCCTTTCGAGGAGTGTGTCGCG 804
Db 482 GGCAACTTCCGCTGGGAGGGCGGTGCGCGCGCGCGATCGGATGCGCTTTCGAGGAGTGTGTCGCG 541
QY 805 TGGCAGCGGGCGCGCTTCCGCGGGGCTGCTTCCAGTGGCGCGCTTCCAGGAGCGCGCG 864
Db 542 TGGCAGCGGGCGCGCTTCCGCGGGGCTGCTTCCAGTGGCGCGCTTCCAGGAGCGCGCG 601
QY 865 CTGCGCATGTTCCAGCGCTTACTGGCCCTGAGCGCGCGAGCGCGCGCGCGCGAGAG 924
Db 602 CTGCGCATGTTCCAGCGCGCTTCTGGCGCTAGAGCGCTGAGCGCGCGCGCGCGCGAGAG 661
QY 925 GTGTTCCGCTTCCCTCAAGCAGAGCTCACTCCGTCGAGCTGGCGCGCGCGCGCG 968
Db 662 GTCTTTCGCTTCCCTCAAGCAGAGCTCACTCCGTCGAGCTGGCGCGCGCGCGCGCG 705

RESULT 15
BU613351 767 bp mRNA linear EST 20-FEB-2003
LOCUS UI-M-EW0-cay-p-20-0-UI.r1 NIH BMAP EW0 Mus musculus cDNA clone
DEFINITION UI-M-EW0-cay-p-20-0-UI 5', mRNA sequence.
ACCESSION BU613351
VERSION BU613351.1 GI:23279566
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .767
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_EW0"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose

```

QY 508 CTGTGTACCGCGACATCAAGCCGAGAACCTGTGTGTTCGACGCGAGTGCCTGCGCG 567
 Db 421 CTGTGTACCGCGACATCAAGCCGAGAACCTGTGTGTTCGACGCGAGTGCCTGCGCG 480
 QY 568 GTAAAGTGGCGGACTTCGACGCGCGCGTGGGCTGCGCGCTCAAGCGGTGAGC 627
 Db 481 GTGAAGTGGTGTACTTCGACGCGCGCGTGGGCTGCGCGCTCAAGCGGTGAGC 540
 QY 628 GGCACCATCCCTTACACGCGCGCTGAGGTGTCCAGCGCGCGCGCGCGCGCTGCGCG 687
 Db 541 GGCACATACCTACACGACCGCGAGGTGTCCAGCGCGCGCGCGCGCGCTGCGCG 600
 QY 688 GTGACACCGCGGTGAGAGTGTGGGCTTCCGCGTGTCTCATCTTCTGCGTGTCTACCGCG 747
 Db 601 GTGACACCGCGGTGAGAGTGTGGGCTTTCGCGTGTCTCATCTTCTGCGTGTCTACCTGCG 660
 QY 748 AACTTCCGCTGGGAGCGGTGTGGGCGCGACGCGCTTCTTCGAGGAGTTCGTGCGCTGG 807
 Db 661 AACTTCCGCTGGGAGGTGTGATCGGCTGCGGATG-CITCTTCGAGGAGTTCGTGCGCTGG 719
 QY 808 CAGCGGGCGCG 818
 Db 720 CAGCGGGCTCG 730

RESULT 13
 BM947152 733 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EH0P-bva-d-11-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
 DEFINITION IMAGE:5689306 5', mRNA sequence.

ACCESSION BM947152
 VERSION BM947152.1
 KEYWORDS GI:19430737
 EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 733)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 733

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5689306"

/tissue_type="whole brain"

/dev_stages="embryo 18.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_EH0P"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is CAGCCACGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institute of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 113 a 216 c 249 g 154 t 1 others

ORIGIN

Query Match 47.3%; Score 603.2; DB 12; Length 733;

Best Local Similarity 89.8%; Pred. No. 6.4e-107;

Matches 658; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

QY 265 ACCAAGCTGAAGAACTTCCTACGCGGAGTACCAACAGCCCTCTCTCCAGCCCC 324

Db 2 ACAAGCTGAAGAACTTCCTGCGTGAAGTACCAACAGCCCTCTCTAGCCCC 61

QY 325 TTCTATCATCAAGTCTTTGAGTGGTCTTTGAGACAGAGACTGCTACGCTTTGCCAG 384

Db 62 TTCTATCATCAAGTCTTTGAGTGGTCTTTGAGACAGAGAGTGTATGCTTTGCCAG 121

QY 385 GAGTACGACCTGCTGGGACCTGTTGACATCATCTCCAGAGTGGGCTCCCTGAG 444

Db 122 GAGTATGACCTGCTGGGACCTGTTGACATCATCTCCAGTGGGCTCCCAAG 181

QY 445 GACACGTGAAGCGCTGTGTGACAGCTGGGCTGGCTGTGACTTCATGCAAGCGCG 504

Db 182 GACACGTGAAGCGCTGTGTGACAGCTGGGCTGGCTGTGACTTCATGCAAGCGCG 241

QY 505 CAGCTGGTGCACCGGACATCAAGCCGAGACGTGCTGTTCGACCGGAGTGGCG 564

Db 242 CAGCTGGTGCACCGTGAATCAAGCCGAGATGCTGTGTTCGACCGGAGTGGCG 301

QY 565 CGCGTAAAGCTGGCGGACTTCGGCATGACGCGCGCGTGGGCTCCCGCTCAAGCGCG 624

Db 302 CGTGTGAAGTGGCTGTGACTTCGGCATGACGCGCGCGTGGGCTCCCGCTGAAGCGCG 361

QY 625 AGCGGCAACCATCCCTTACAGCGCGCTGAGTGTGCCAGGCGCGCGCGCGCGCG 684

Db 362 AGCGGCACTATACCGTACAGCACCGCGGCTGTGCCAGGCTGGCGCGCGCGCGGCTTC 421

QY 685 GCGGTGACACGCGCGTGGAGTGTGGGCTTGGCGCTTCATCTTCTGGGCTCTACC 744

Db 422 GCGGTGACACGCGCGTGGAGTGTGGGCTTGGCGCTTTCATCTTCTGGGCTCTACT 481

QY 745 GGCAACTTCCCGTGGGAGGCGCGTCCGCGCGCGCGCGCGCGCTTCCTCGAGGAGTTCGTGCG 804

Db 482 GGCAACTTCCCGTGGGAGGCTGCACTCGGTGCGGATGCCCTTCCTCGAGGAGTTCGTGCG 541

QY 805 TGGCAGCGGGCGCGCTGCGCGGCTGCTTTCGAGTGGCGCGCGCTTCACCGAGCGCG 864

Db 542 TGGCAGCGGGCTGCGCGGCTGCTTCATCGCAGTGGCGCGCGCTTTTACGAGCGCTGCG 601

QY 865 CTGCGCATGTTCACAGCGCTTACTGCGCTTGAGCGCGCGCGCGCGCGCGCAAGAG 924

Db 602 CTGCGTATGTTCACAGCGCTTCTGCGCTTAGAGCGCGCGCGCGCGCGCAAGAG 661

QY 925 GTGTTCCGCTTCCCTCAAGCAGCAGCTCAAGTCCAGTGGCGCGCGCGCGCGCTTCGACCGCG 984

Db 662 GTCTTTGCTTCCCTCAAGCAGCAGCTCAAGTCCAGTGGCGCGCGCGCGCTTCGACCGCG 720

QY 985 GCGGCAAGCGCG 997

Db 721 GCAGCAAGCGCG 733

RESULT 14

BM947988

LOCUS

DEFINITION

IMAGE:5703871 5', mRNA sequence.

ACCESSION

BQ179888

UI-M-EW0-bww-c-08-0-UI.r1 NIH_BMAP_EW0 Mus musculus cDNA clone

IMAGE:5703871 5', mRNA sequence.

BQ179888

706 bp mRNA linear

EST 30-APR-2002

```
JOURNAL
COMMENT

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgraphs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP).

Seq primer: pYX-5.

Location/Qualifiers
1. 730
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 5721292"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP F10"
/notes="Organ: Brain; Vector: pYX- Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CACCACGAC. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT      139 a      197 c      241 g      153 t
ORIGIN

Query Match      47.4%; Score 603.8; DB 13; Length 730;
Best Local Similarity 90.0%; Pred. No. 4.9e-107;
Matches 658; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

Qy 88 CCCCTTCTCACTGAAGACATGCGGCGCTGACTTCGCGACACTGGCGCGGACGACGTC 147
Db 1 CCCCTCTCACAAGACATGCAAGCGCTGACCTACCGACACTGGCTGGCGGCGATGTT 60

Qy 148 ACCAGCACTACGAACTAGTCCGGAGCTGGGCAAGGACCTATGGGAAGTTGACCTG 207
Db 61 ACCAAGCACTATGAGCTGTGCGGGAGCTGGGCAAGGACCTACGGAAGTTGACCTG 120

Qy 208 GTGGCTTACAAGGGCACAGGCACAAAATGSCACTGAAGTTTGTGAACAAGACAAACC 267
Db 121 GTGGCTTACAAGGGCACAGGCACAAAATGSCACTGAAGTTTGTGAACAAGTAAGACA 180

Qy 268 AAGCTGAAGAACTTCTTACCGGAGGTGAGCATACCAACAGCTCTCTCCAGCCCTTC 327
Db 181 AAGCTGAAGAACTTCTTACCGGAGGTGAGCATACCAACAGCTCTCTAGCCCTTC 240

Qy 328 ATCATCAAGCTCTTGAAGCTGCTTTCAGACAGAGGACTGCTACGCTTTGCCAGGAG 387
Db 241 ATCATCAAGCTCTTGAAGCTGCTTTCAGACAGAGGAGTCTATGCTTTGCCAGGAG 300

Qy 388 TACGCACTCTGCGGGACCTGTTTGACATCATCTCCCTCCAGGTGGGGCTCCCTGAGGAC 447
Db 301 TATGCACTCTGCGGGACCTGTTTGACATCATCTCCCTCAGGTGGGGCTCCCAAGGAC 360

Qy 448 ACGGTGAAGCGCTGTGTGACAGCACTGCGCTGGCGCTGGAAGTCTATGACGGCGGAG 507
Db 361 ACGGTGAAGCGCTGTGTGACAGCACTGCGCTGGCGCTGGAAGTCTATGACAGGCGAG 420

Query Match      47.6%; Score 606.8; DB 12; Length 746;
Best Local Similarity 90.0%; Pred. No. 1.3e-107;
Matches 672; Conservative 0; Mismatches 73; Indels 2; Gaps 2;

Qy 218 AGGCAAGGACACAAAATGCACTGAAGTTGTGAACAAGACAAACCAAGCTGAAGA 277
Db 2 AGGCAAGGACACAAAATGCACTGAAGTTGTGAACAAGTAAGTAAGCAAAAGCTGAAGA 61

Qy 278 ACTTCTTACGGGAGGTGAGCATCAACCAAGCCTCTCTCCAGGCCCTTTCATCATCAAG 337
Db 62 ACTTCTTACGGGAGGTGAGCATCAACCAAGCCTCTCTAGCCCTTTCATCATCAAG 121

Qy 338 TCTTTGACGTGCTTTTGAAGACAGAGACTGCTACGCTTTTGGCCAGGAGTACGACCTG 397
Db 122 TCTTTGACGTGCTTTGACAGACCGAGAGTCTATGCTTTTGGCCAGGAGTATGACCTG 181

Qy 398 CTGGGACCTGTTTGAAGCATATCCCTCCAGGTTGGGCTCCCTGAGGACACAGTGAAGC 457
Db 182 CTGGGACCTGTTTGAAGCATATCCCTCTCAGTTGGGCTCCCAAGGACACAGTGAAGC 241

Qy 458 GCTGTGTGACAGCTGGGCTGCGCTGAGCTTTCATGACGGGCGGAGCTGTGTGCAAC 517
Db 242 GCTGTGTGACAGCTGGGCTGCGCTGAGCTTTCATGACAGGACGAGCTGTGTGCAAC 301

Qy 518 GCGACATCAAGCCGAGAACGCTGCTGTTTCGACCGGAGTGGCGCGGTAAAGCTGG 577
Db 302 GTGACATCAAGCCGAGAACGCTGCTGTTTGAACCGGTGAGTGGCGCGGTGAAAGCTGG 361

Qy 578 CCGACTTCGGCATGACGGCGCGGTGGGCTGCGCGGTCAAGCGGTGAGCGGACCATCC 637
Db 362 CTGACTTCGGCATGACGGCGCGGTGGGCTGCGCGGTGAAAGCGGTGAGCGGACATAC 421

Qy 638 CTTACACGGCCCTGAGGTGTCAGGCGGCGCGCGCGGCTGCGGCTGCGGTGACACGG 697
Db 422 CGTACACAGCACCCGAGGTGTCAGGCTGGCGCGCGGAGTTCGCGGTGACACAGG 481

Qy 698 GCGTGGAGCTGCGGCTTCCGCGGTGCTCATCTTCTGCTGCTCACCGGCACTCCCGT 757
Db 482 GCGTGGAGCTGCGGCTTTCGCGGTGCTCATCTTCTGCTGCTCACGCGCACTTCNGT 541

Qy 758 GGGAGGCGGCTCGGGCGCGGCTTCTTTCGAGGAGTTCGTGCGCTGGCAGCGGGGCC 817
Db 542 GGGAGGCTGATCGGGTGGGATGCTTTCGAGGAGTTCGTGCGCTGGCAGCGGGGTC 601

Qy 818 GCCTGCGGGGCTGCTTTCGAGTGGCGGCTTTCACCGACCGCGCTGCGCATGTTCC 877
Db 602 GCCTGCGGGGCTG-CATCGAGTGGCGCGCTTTCACCGAGCTGCGCTGCGTATGTTCC 660

Qy 878 AGCGCTTACTGCGCTGAGCGCGGCGCGCGGCGGCGGCGGAGGAGTGTTCGCTTCC 937
Db 661 AGCGCTTCTGCGCTGAGCGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719

Qy 938 TCAAGCAGAGCTACGCTCCGAGCTGC 964
Db 720 TCAAGCAGAGCTACATCCGAGCTGC 746

RESULT 12
BU701617
LOCUS
DEFINITION
IMAGE: 5721292 5', mRNA sequence.
ACCESSION
BU701617
VERSION
BU701617.1
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 730)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
```


CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

Seq primer: pYX-5
Location/Qualifiers
1..719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5703632"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP EWO"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemm Chin, Ph.D., program coordinator."

BASE COUNT

117 a 202 g 158 t
Query Match 47.6%; Score 607; DB 13; Length 719;
Best Local Similarity 90.3%; Pred. No. 1.2e-107;
Matches 649; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

247 TTTCTGAACAGAGCAAAACCAAGCTCAAGACTTCTTACGGAGGTGAGCATCACCAAC 306
Db 1 TTTCTGAATAGATGACACAAAGCTGAGACTTCTGCGTGAAGTGAGCATCACCAAC 60
307 AGCTCTCTCCAGCGCTTCATCATCAAGTCTTTGACGGTCTTTGACAGAGGAC 366
Db 61 AGCTCTGCTCTAGCCCTTCATCATCAAGTCTTTGACGGTCTTTGACAGAGGAG 120
367 TGCTACGCTTTTGCCAGGAGTACGACCTCTGCGGACCTGTTGACATCATCCCTCC 426
Db 121 TGCTATGCTTTTGCCAGGAGTATGACCTCTGCGGACCTGTTGACATCATCCCTCC 180
427 CAGTGGGGTCTCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCTGGCGCTG 486
Db 181 CAGTGGGGTCTCCAGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCTGGCGCTG 240
487 GACTTCATGACAGCGCGAGCTGTGACCGCGACATCAAGCCCGAGAGACGTCTGCTG 546
Db 241 GACTTCATGACAGCAGCGAGCTGTGACCGCGACATCAAGCCCGAGAGATGTCTGCTG 300
547 TTCACCGCGAGTCCCGCGTAAAGCTGSCCGACTTCGGCATGACGCCCGCTGGGC 606
Db 301 TTTGACCGTGAAGTCCCGCGTGAAGCTGTGACCTGTGACATGACGCCCGCTGGGC 360
607 TGCCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACACGGCGCTGAGGTGTGCCAGGCG 666
Db 361 TGCCGCTGAAGCGCGTGAAGCGGACCTATACCGTACACAGCACCAGAGGTGTGCCAGGCT 420
667 GGCCGCCCGCAGCGGCTGGCGGTGGACACGGCGGTGACGTGTGGGCTTTCCGCGTCTC 726
Db 421 GGCCGCCCGCATGCTTCCGCGGTGGACACGGCGGTGACGTGTGGGCTTTGGCGTCTC 480

QY 727 ATCTTCTGCTGCTACCGCAACTTCCGTGGAGGCGGCGTCCGGCGCGAGCCCTTC 786
Db 481 ATCTTCTGCTGCTACTGGCAACTTCCGTGGAGGCGTCCGTGGCGATGCGCTTC 540
QY 787 TTCAGAGATTGCTGCGTGGCAGCGGCGCTGCGGGGCTGCTTCCGAGTGGCGC 846
Db 541 TTCAGAGATTGCTGCGTGGCAGCGGCGCTGCGGGGCTGCTTCCGAGTGGCGC 600
QY 847 CGTTCACCGAGCCCGCTGCGCATGTTCCAGGCTTACTTGGCCCTGGAGCCGAGCGC 906
Db 601 CGTTCACCGAGCCCTGCGTATGTTCCAGGCTTCTGGCGCTAGACCTGAGCGG 660
QY 907 CGCGGCCAGCAGGAGGTTCGCTTCTCTCAAGCAGAGCTCAGTCCGAGTGGG 965
Db 661 CGTGGGCCCGCCAGGAGGTCTTTCGCTTCTCTCAAGCAGAGCTCAGTCCGAGTGGG 719

RESULT 11

BM944369
LOCUS
DEFINITION
UI-M-EH0p-bvq-j-11-0-UI.r1 NIH BMAP_EH0p Mus musculus cDNA clone
IMAGE:5695594 5', mRNA sequence.
BM944369
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 746)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
1..746
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695594"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_EH0p"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCAGCAGC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemm Chin, Ph.D., program coordinator."

BASE COUNT
127 a 206 c 251 g 161 t
ORIGIN

```

QY 988 GCGAAGCCCCCGGGGACCGCCCGCCCGCCGCGCCACCTGCGCTCGAGCGCCTGGG 1047
Db 661 CGCAAGCGCGCTGNGAGACCGGCTGC-----CAGGGTCNCTGCGCCTTGAGGCTCCCGG 714
QY 1048 CCGCTCAAGCGGAGCGGTCTGACCGAGAGCGGACGCGCTCCG 1091
Db 715 GCATCAAGCGCACTGTGCTCACCGAGAGTGCGACGGCTCGG 758

RESULT 9
CA450675 833 bp mRNA linear EST 26-NOV-2002
LOCUS UI-M-FW0-cbp-g-09-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
DEFINITION IMAGE: 6812674 5', mRNA sequence.
ACCESSION CA450675
VERSION CA450675.1 GI:24815095
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library preparation: Dr. Jim Lin, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

The following repetitive elements were found in this cDNA
sequence: 39-77, >G-rich#Low_complexity
Seq primer: pYX-5.

FEATURES             Location/Qualifiers
     source            1..833
     organism="Mus musculus"
     mol_type="mRNA"
     strain="C57BL/6"
     db_xref="taxon:10090"
     clone="IMAGE: 6812674"
     tissue_type="whole brain"
     dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
     lab_host="DH10B (T1 phage resistant)"
     clone_lib="NIH_BMAP_FW0"
     note="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I;
     Site:2: Not I; The library was constructed according
     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
     1996. Denatured RNA was size fractionated on a 1% agarose
     gel. First strand cDNA synthesis was primed with oligo-dT
     primer containing a Not I site. Double strand cDNA was
     size selected according to mRNA size fraction, ligated
     with EcoR I adaptor, digested with NotI and then cloned
     directionally into pYX-Asc vector. The library tag
     sequence located between the Not I site and the polyA tail
     is ACGGAGACAG. This library was created for the University
     of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
     Developing Mouse Nervous System', supported by National
     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
     program coordinator."

BASE COUNT 154 a 247 c 273 g 159 t
ORIGIN

Query Match 47.8%; Score 609.2; DB 14; Length 833;
Best Local Similarity 88.8%; Pred. No. 4.5e-108;
Matches 659; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```

```

QY 1 ATGAGCTGGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
Db 92 ATGAGCTGGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 151
QY 61 ACTCCCTCTGGGCTGCTGGCGGTGTCGCCCTTCTCACTGAAGACATGACGGCCCTGACT 120
Db 152 GCGGCCCTCTGACAGGTGACAGGTGTCGCCCTTCTCAAGAGACATGACAGCCCTGACC 211
QY 121 CTCGCGACACTGGCGCCGACGAGCGTCAACAAGCACTACGAACCTAGTCCGGAGCTGGGC 180
Db 212 CTACGCACTGCTGGCTGCCAGCATGTTTACCAAGCACTATGAGCTTGTGGGAGCTGGGC 271
QY 181 AAAGGCACTTATGGGAAGGTGTGACCTGGTGGTCTACAAGGCAAGGCAAAAAATGGCA 240
Db 272 AAAGGCACTTATGGGAAGGTGTGACCTGGTGGTCTACAAGGCAAGGCAAAAAATGGCC 331
QY 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATC 300
Db 332 CTGAATTTGTGAATAAGAGTAAAGCAAGCTGAAGAACTTCTTGGCTGAGGTGAGCATC 391
QY 301 ACCAAGCCCTCTCTCCAGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA 360
Db 392 ACCAAGCCCTCTCTCTAGCCCTTTCATCAAGGTCTTTGACGTGGTCTTTGAGACC 451
QY 361 GAGGACTGTACTGTTTGGCCAGGAGTACGCACCTGCTGGGACCTGTTTGACATCATC 420
Db 452 GAGGAGTGCTATGTTTGGCCAGGAGTATGCACCTGCTGGGACCTGTTTGACATCATC 511
QY 421 CCTCCCGAGTGGGGCTCCCTGAGGACAAGGCTGAAGCGCTGTGTGACGAGCTGGGCTG 480
Db 512 CTCTCTCAGGTGGGGCTCCAGAGAGACAAGTGAAGCGCTGTGTGACGAGCTGGGCTG 571
QY 481 GCGCTGAGCTTCATGCAAGCGGCGGCGCTGTGTGCAACCGACATCAAGCCCGAGAACGTG 540
Db 572 GCGCTGAGCTTCATGCAAGCAGCAGCAGCAGTGTGTGCAACCGCTGACATCAAGCCCGAGAACGTG 631
QY 541 CTGCTGTTCGACCCGAGTGGCGCGCTAAAGCTGACCGCTTCGCGATGACGAGCTGGCG 600
Db 632 CTGCTGTTCGACCCGAGTGGCGCGCTAAAGCTGACCGCTTCGCGATGACGAGCTGGCG 691
QY 601 GTGGGCTGCCGCGTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGCGCTGAGGTGTC 660
Db 692 GTGGGCTGCCGCTGAAGCGCGTGAAGCGGACCACTATACCGTACACAGCACCGGAGTGTG 751
QY 661 CAGCGCGCGCGCGCGCGCGCTGGCGTGTGAACAAGCGCGTGGACGTGTGGGCGCTTCGCG 720
Db 752 CCAGGCTGGCGCGCGCGATGCGCTTCGCGTGGACACCGGCGCGTGGACGTGTGGGCGCTTGGC 811
QY 721 GTGCTCATCTTCTGCGTGCTCA 742
Db 812 GTGCTCATCTTCTGTGTGCTCA 833

RESULT 10
LOCUS BQ745423 719 bp mRNA linear EST 17-JUL-2002
DEFINITION UI-M-EW0-bwv-i-09-0-UI.r1 NIH_BMAP_EW0 Mus musculus cDNA clone
IMAGE:5703632 5', mRNA sequence.
ACCESSION BQ745423
VERSION BQ745423
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

```


oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GTCGCGGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 108 a 243 c 275 g 158 t 2 others
 ORIGIN
 Query Match 48.6%; Score 619.2; DB 13; Length 786;
 Best Local Similarity 87.8%; Pred. No. 5.2e-110;
 Matches 688; Conservative 0; Mismatches 90; Indels 6; Gaps 1;
 QY 336 GGTCTTTGACGTGGTCTTTGAGACAGAGGACTGCTACGCTTTTGCACAGGAGTAGCGACC 395
 DB 1 GGTCTTTGACGTGGTCTTTGAGACAGGAGTGTCTATGCTTTTGCACAGGAGTAGCGACC 60
 QY 396 TGCTGGGACCTGTTTGACATCATCCCTCCAGGTGGGCTCCCTGAGGACACGGTGAA 455
 DB 61 TGCTGGGACCTGTTTGACATCATCCCTCCAGGTGGGCTCCAGAGGACACGGTGAA 120
 QY 456 GCGCTGTGTGACAGCTGGGCTGGCGCTGAGCTTCATGACAGGCGGCGAGCTGTGCA 515
 DB 121 GCGCTGTGTGACAGCTGGGCTGGCGCTGAGCTTCATGACAGGCGGCGAGCTGTGCA 180
 QY 516 CCGGACATCAAGCCGAGAGACGTGCTGCTGTGACCGCGAGTGCCTCCGCTAAAGCT 575
 DB 181 CCGTGACATCAAGCCGAGATGTGCTGCTGTTGACCGTAGTGCCCGCTGTGAAGCT 240
 QY 576 GGCGACTTCGGCATGACGCGCGGCTGGCTGGCGCTCAAGCGCGTGAAGCGGACCAT 635
 DB 241 GGCTGACTTCGGCATGACGCGCGGCTGGCTGGCGCTGAAGCGCGTGAAGCGGACCAT 300
 QY 636 CCCTTACAGCGGCTGAGTGTGTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 695
 DB 301 ACCCTTACAGCAGCGGAGTGTGCGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 QY 696 GGGCGTGACGTGTGGGCTTTCGGGCTGCTCATCTTCTGCGTGTCTACCGCACTTCCC 755
 DB 361 GGGGTGACGTGTGGGCTTTCGGGCTGCTCATCTTCTGCGTGTCTACCGCACTTCCC 420
 QY 756 GTGGAGCGCGGCTGCGGCGCGGCGGCTTCTTCAGAGAGTTCGTGCGCTGCGAGCGGG 815
 DB 421 GTGGAGGCTGCTATCGGCTGCGGATGCTTCTTCAGAGAGTTCGTGCGCTGCGAGCGGG 480
 QY 816 CCGCTGCGGCGGCTGCTTTCAGAGTGTGCGGCGGCTTTCAGAGAGTTCGTGCGCTGCGAGCGGG 875
 DB 481 TCGGCTGCGGCGGCTGCTTTCAGAGTGTGCGGCGGCTTTCAGAGAGTTCGTGCGCTGCGAGCGGG 540
 QY 876 CCAGCGCTTACTGCGCTTTCAGAGTGTGCGGCGGCTTTCAGAGAGTTCGTGCGCTT 935
 DB 541 CCAGCGCTTACTGCGCTTTCAGAGTGTGCGGCGGCTTTCAGAGAGTTCGTGCGCTT 600
 QY 936 CCTCAAGACAGAGCTCAGTCCGAGTGTGCGGCGGCGGCTTTCAGAGAGTTCGTGCGCTT 995
 DB 501 CCTCAAGACAGAGCTCAGTCCGAGTGTGCGGCGGCGGCTTTCAGAGAGTTCGTGCGCTT 660
 QY 996 CCGCGGAGCGCGCGCGCGCGCGGCGGCTTTCAGAGAGTTCGTGCGCTTTCAGAGAGTTCGTGCGCTT 1055
 DB 661 GCCTGGGACCGCGCTGCGGCGGCGCGGCGGCTTTCAGAGAGTTCGTGCGCTTTCAGAGAGTTCGTGCGCTT 714
 QY 1056 GCGGACGCTGTGACCGAGAGCGGCGGCGGCTTTCAGAGAGTTCGTGCGCTTTCAGAGAGTTCGTGCGCTT 1115
 DB 715 GCGGACGCTGTGACCGAGAGCGGCGGCGGCTTTCAGAGAGTTCGTGCGCTTTCAGAGAGTTCGTGCGCTT 774
 QY 1116 GGTG 1119
 DB 775 CGTG 778

RESULT 7
 CA324262
 LOCUS
 DEFINITION
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CA324262 742 bp mRNA linear EST 26-NOV-2002
 UI-M-PYO-cco-1-22-0-UI.r1 NIH BMAP_FYO Mus musculus cDNA clone
 IMAGE: 6822407 5', mRNA sequence.
 CA324262
 EST
 CA324262.1 GI:24542360
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers
 1. 742
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6822407"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpf"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP_FYO"
 /notes="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to BMAP, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGGAGAGAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institutes of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."
 BASE COUNT 128 a 199 c 255 g 159 t 1 others
 ORIGIN

Query Match 48.4%; Score 617.2; DB 14; Length 742;
 Best Local Similarity 89.9%; Pred. No. 1.3e-109;
 Matches 661; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 171 GGAGCTGGGCAAGGACCTTATGGAGGTTGACCTGGTGTGTACAGGGCACAGGCAC 230
 DB 8 GGAGCTGGGCAAGGACCTTATGGAGGTTGACCTGGTGTGTACAGGGCACAGGCAC 67
 QY 231 AAAATGGCACTGAAGTTTGTGAACAGCAAAACCAAGCTCAAGAACTTCTACGGGA 290
 DB 68 TAAATGGCCCTGAAATTTGTGAATAGAGTAAAGACAAAGCTCAAGAACTTCTGCTGA 127
 QY 291 GGTGAGCATCAACCAAGAGCTTCTCCAGCCCTTCATCATCAAGCTTTTACAGCTGGT 350
 DB 128 GGTGAGCATCAACCAAGAGCTTCTCCAGCCCTTCATCATCAAGCTTTTACAGCTGGT 187

QY 481 CGGTGACCTTATGACAGCGCGGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGT 540
 |||||
 Db 497 CGCTGGACTTATGACACAGCAGGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGT 556
 |||||
 QY 541 CTGCTGTTTCGACCGGAGTGCAGCGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGC 600
 |||||
 Db 557 CTGCTGTTTACCGGTAGTGCAGCGCGTGAAGCTGGCTGACTTCGGCATCAGCGCGC 616
 |||||
 QY 601 GTGGGCTGCCCGTCAAGCGGCTGAGCGGCACCTCCCTTACAGCGGCGCTGAGTGTGC 660
 |||||
 Db 617 GTGGGCTGCCGTGTAAGCGGCTGAGCGGCACCTATACCGTACACAGCACCGGAGTGTGC 676
 |||||
 QY 661 CAGCGGCGCGCGCGAGCGCTGGCGGTGACACAGCGCGCTGGAGCTGTGGGCTTCGGC 720
 |||||
 Db 677 CAGCTGCGCGCGCGAGTGGCTTCGCGGTGACACAGCGCGCTGGAGCTGTGGGCTTCGGC 736
 |||||
 QY 721 GTGCTCATCTTCTGCGTCTCACCGG 746
 |||||
 Db 737 GTGCTCATCTTCTGCTCACTGG 762
 |||||

RESULT 4
 BM964264 746 bp mRNA linear EST 18-MAR-2002
 LOCUS UI-M-EQ0-bvu-m-02-0-UI.r1 NIH_BMAP_EQ0 Mus musculus cDNA clone
 DEFINITION IMAGE:5697193 5', mRNA sequence.
 ACCESSION BM964264
 VERSION BM964264.1 GI:19547684
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 746)
 NIH-MGC <http://hgnc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..746
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5697193"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EQ0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCACGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous

FEATURES

source

System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 139 a 209 c 244 g 154 t
 ORIGIN
 Query Match 48.6%; Score 619.6; DB 12; Length 746;
 Best Local Similarity 89.4%; Pred. No. 4.3e-110;
 Matches 667; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 2 TGACGCTGGGCTGCCACAGAGCTGAGCGCGCCGCTCCCTGACCTGCTGTGGGCGGGGA 61
 |||||
 Db 1 TGACGCTGGGCTGCCCTGAGCCTGAACCGCTCACCTCCCTGCTGTGTGGCGGGGG 60
 |||||
 QY 62 CTGCCCCCTGGGCTGTGGTGGCGGTGCCCCCTTCTACTGAAGACATGACAGCCCTGACTC 121
 |||||
 Db 61 CCGCCCCCTTACCAGGTGCAGGTGTGCCCTCTCTACAGAAGACATGCAAGCGGTGACCC 120
 |||||
 QY 122 TCGCACACTGGCGCCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGAGAGCTGGGCA 181
 |||||
 Db 121 TACGCACACTGGCTGCCAGCGATGTACCAAGCACTATGAGCTTGTCCGGGAGCTGGGCA 180
 |||||
 QY 182 AAGCACCTATGGAGAGGTTGACCTGTGTGTCTACAAGGGCACAGGCACAAAATGGCAC 241
 |||||
 Db 181 AAGGGACCTACGGGAAGGTGACCTGTGTGTCTACAAGGGCACAGGCACCTAAAATGGCCC 240
 |||||
 QY 242 TGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATCA 301
 |||||
 Db 241 TGAATTTGTGAATAAGAGTAAGACAAAGCTGAAGAACTTTCTGCGTGAAGTGAAGCATCA 300
 |||||
 QY 302 CCACAGCCTCTCTCCAGCCCTTCTATCATCAAGGTCTTTGAGTGGTCTTTGAGACAG 361
 |||||
 Db 301 CCAACAGCCTGTCTGTAGCCCCCTTCAATCATCAAGGTCTTTTGAAGTGTCTTTGAGACCG 360
 |||||
 QY 362 AGGACTCTACGTCTTTGCCAGGAGTACGACCTGTCTGGGGACCTGTTTGACATCATCC 421
 |||||
 Db 361 AGGAGTCTATGCTTTTGGCCAGGAGTATGCACTGTCTGGGGACCTGTTTGACATCATCC 420
 |||||
 QY 422 CTCGCCAGGTGGGGCTCCCTGAGACACAGCTGAAGCGCTGTGTGACAGAGCTGGGCTGG 481
 |||||
 Db 421 CTCCTCAGGTGGGGCTCCCGAGAGACACGGTGAAGCGCTGTGTGACAGAGCTGGGGTGG 480
 |||||
 QY 482 CGCTGGACTTTCATCACGGCGGCGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGC 541
 |||||
 Db 481 CGCTGGACTTTCATGCAAGAGGAGCGGTGTGTGACCGTGACATCAAGCCCGAGAACTGTC 540
 |||||
 QY 542 TGTGTTTCGACCGGAGTGGCGCGCTAAAGCTGTGGCGACTTCGGCATGACGCGCGCG 601
 |||||
 Db 541 TGTGTTTTCACCGTGAAGTGGCGCGCTGAAGCTGTGACTTCGGCATGACGCGCGCG 600
 |||||
 QY 602 TGGGCTGGCGGCTCAAGCGGTGAGCGGCACCATCCCTTACAGCGGCGCTGAGGTGGCC 661
 |||||
 Db 601 TGGGCTGGCGGTTGAAGCGGTGAGCGGCACCTATACCGTACACAGCACCGGAGTGTGCC 660
 |||||
 QY 662 AGCGGGCGCGCGCGAGCGGCTGGCGGTGGACACAGCGGCGTGGAGCTGTGGGCTTCGGCG 721
 |||||
 Db 661 AGGTGCGCGCGCGATGGCTTCGCGGTGGACACAGCGGCGTGGAGCTGTGGGCTTTGGCG 720
 |||||
 QY 722 TGCTCATCTTCTGCGTGTCTCACCGGC 747
 |||||
 Db 721 TGCTCATCTTCTGCTGCTCACTGGC 746
 |||||
 RESULT 5
 BU611098 762 bp mRNA linear EST 20-FEB-2003
 LOCUS UI-M-PC0-cas-f-08-0-UI.r1 NIH_BMAP_PC0 Mus musculus cDNA clone
 DEFINITION UI-M-PC0-cas-f-08-0-UI 5', mRNA sequence.
 ACCESSION BU611098
 VERSION BU611098.1 GI:23277313
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

```
seq format="gff3"
location/Qualifiers
1..762
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UT-M-PCO-cas-b-20-0-UT"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resis"
/clone_lib="NIB MAP ECO"
```

	(NINT)	HENRI CHIL,	FIB.,	program	coordin
BASE COUNT	145 a	274 c	248 d	154 f	1 others

Query Match 48 68: score 620 3: DB 13: Length 762:

Query Match	48.6%	Score 620.2;	DB 13;	Length 762;
Best Local Similarity	89.4%	Pred No 3	3e-110.	

Best Local Similarity 89.4%; Pred. NO. 3.3e-110;
Matches 667; Conservative 0; Mismatches 78; Indels 0; Cons 0.

Qy	1	ATGAGCGTGGCGTCCGAGAGCTGAGCCGCGCCGCTCCCTGACCTCGTGTGGGCGCGGG	60
Db	17	ATGAGCGTGGCGTCCGCTGAGCTGAAACGCTCCACTCCCTGCTTCTGTGGGCGCGGG	76
Qy	61	ACTGCCCTGGGCTGTGTGCGGTGTGCCCTTCTCACTGAAGACATGCAAGGCCCTGACT	120
Db	77	GCGGCCCTGTACCGAGTGCAGGTGTGCCCTCTCTACAGAGACATGCAAGCGCTGACC	136
Qy	121	CTCCGCACACTGGCGCGCAGCGAGCTCAACGAGCCTACGAACTGAGTCCCGGAGCTGGGC	180
Db	137	CTACGCACACTGGCTGCGCAGCGATGTTACCAAGCACCTATGAGCTTGTGCGGAGCTGGC	196
Qy	181	AAAGGCACCTATGGGAAGGTTGACCTGTGTCTTACAAGGGCAGAGCACAAAAATGC	240
Db	197	AAAGGACCTACCGGAAGGTGACCTGTGTGCTTAAAGGGCAGAGCACTAAATGTC	256
Qy	241	CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC	300
Db	257	CTGAAATTTGTGAATGAAGTAAGACAAAGCTGAAGAACTTCTGCGTGAAGTGAGCATC	316
Qy	301	ACCAACAGCTCTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA	360
Db	317	ACCAACAGCTCTGCTGTAGCCCTTCATCATCAAGGTCTTTGACGTGGTCTTCGAGACC	376
Qy	361	GAGGACTGTCTCTTTGCCCCAGGAGTACGCACCTGCTGGGAGACCTGTTTGACATCATC	420
Db	377	GAGGAGTGTATGTCTTTGCCAAGGAGTATGCACCTGCTGGGACCTGTTTGACATCATC	436
Qy	421	CTTCCCGAGGTGGGCTCCCTGAGGACAACGGTGAAGCGCTGTGTGACAGACGTGGGCGTG	480
Db	437	CTTCTCTCAGGTGGGCTCCAGAGGACAACGGTGAAGCGCTGTGTGACAGACGTGGGCGTG	496

Seq primer: pYX-5.
 Location/Qualifiers
 1. .804
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6402026"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP FCO"
 /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; the library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fractionation, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 142 a 221 c 272 g 169 t
 ORIGIN
 Query Match 52.6%; Score 671.2; DB 13; Length 804;
 Best Local Similarity 89.7%; Pred. No. 4.4e-120;
 Matches 721; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 117 GACTCTCCGCACACTGCGCCGCGAGCGAGTCACCAAGCACTAGCACTAGCTGCGGAGCT 176
 Db 1 GACCCTACGCACACTGCTGCCAGGATGTTACCAAGCACTAGCTGCGGAGCT 60
 QY 177 GGGCAAGGCACCTATGGGAAGTTGACCTGGTGTCTCAAGGGCACAGGCAAAAT 236
 Db 61 GGGCAAGGCACCTACGGGAAGTTCGACCTGGTGGCTTACAAGGGCACAGGCAAAAT 120
 QY 237 GGCACCTGAAGTTGTGAACAGAGCAAAACCAAGCTGAAGCACTTCTACGGAGGTGAG 296
 Db 121 GGCCTGAATTTGTGAATGAAGTAGTAAGCAAGCTGAAGCACTTCTGCGTGAAGTGA 180
 QY 297 CATCACCAACAGGCTCTCCCTCCAGCCCTTCATCATCAAGGCTTTTGACGTGCTTTGA 356
 Db 181 CATCACCAACAGGCTCTGCTAGCCCTTCATCATCAAGGCTTTTGACGTGCTTCGA 240
 QY 357 GACAGAGGACTGCTACTCTTTGCCAGGAGTAGCGACCTGCTGGGACCTGTTTGACAT 416
 Db 241 GACCGAGGAGTGTATGTTCTTTCCCAAGGAGTAGCACCTGCTGGGACCTGTTTGACAT 300
 QY 417 CATCCCTCCCAAGTGGGGCTCCCTGAGGACACCGTGAAGCGTGTGTGACAGAGCTGGG 476
 Db 301 CATCCCTCCTCAGTGGGCTCCAGAGGACACCGTGAAGCGTGTGTGACAGAGCTGGG 360
 QY 477 CTGGGCGCTGGAATTCATGCAAGGGGGAGCTGGTGTGACCGGACATCAAGCCCGAGAA 536
 Db 361 GCTGGCGCTGGAATTCATGCAAGGAGGAGCTGGTGTGACCGTGAATCAAGCCCGAGAA 420
 QY 537 CGTGTGCTCTTCGACCGGAGTGCCTCCGCTAAAGCTGGCGACTTCGGCATGACGG 596
 Db 421 TGTGCTGCTGTTTGAACCGTGAAGTGCCTCCGCTGAAGCTGGCTGACTTCGGCATGACGG 480
 QY 597 CCGCGTGGGCTCGCGCTCAAGGCGCTGAGCGGACCATCCCTACACGGCGCTGAGGT 656
 Db 481 CGCGTGGGCTCGCGTGTGAAGCGCTGAGCGGACCTATACCGTACACGACCCGAGGT 540
 QY 657 GTGCCAGGCGGCGCGCGGCTGGCGGCTGGCGGTGGACCGGGGTGGACGTGTGGGCTT 716
 Db 541 GTGCCAGGCTGGCGGCGCGGATGGCTTCGCGGTGGACACGCGGCTGGACGTGTGGGCTT 600

QY 717 CGCGCTGCTCATCTTCTGGTGTCTCACCGGCAACTTCCCGTGGGAGGCGGCGCTCGGCGCG 776
 Db 601 TGGCGTGTCTCATCTTCTGTGTGCTCACTGGCAACTTCCCGTGGGAGGCTGCATCGGCTGC 660
 QY 777 CGAGCGCTTCTTCGAGGAGTTGTGTGCTGGCAGCGGCGCGCTGCGGGGCTGCTTC 836
 Db 661 GGATGCTTCTTCGAGGAGTTGTGTGCTGGCAGCGGCGGCTGCGCTGCGGGGCTGCCATC 720
 QY 837 CGAGTGGCGCGCTTCACCGAGCCCGCTGGCGCATGTTCCAGCGCTTACTGSCCTTGA 896
 Db 721 CGAGTGGCGCGCTTACCGAGCTGCTGTGTGTAATGTTCCAGCGCTTCTGCGGCTAGA 780
 QY 897 GCCGAGCGCGCGCGCCAGCCAA 920
 Db 781 GCGTAGCGCGTGGCGCCGCA 804
 RESULT 2
 EG683070
 LOCUS 602651746F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:4761754 5', mRNA linear EST 01-MAY-2001
 DEFINITION mRNA sequence.
 ACCESSION EG683070
 VERSION EG683070.1 GI:13914467
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 888)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCMI614 row: 1 column: 11
 High quality sequence stop: 682.
 FEATURES
 Location/Qualifiers
 1. .888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4761754"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 47"
 /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally inserted into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."
 BASE COUNT 128 a 312 c 325 g 123 t
 ORIGIN
 Query Match 49.5%; Score 630.6; DB 10; Length 888;
 Best Local Similarity 91.4%; Pred. No. 3.3e-112;
 Matches 736; Conservative 0; Mismatches 59; Indels 10; Gaps 6;
 QY 454 AAGCGTGTGTGACGAGCTGGCGCTGGGCTGGGACTTCATCAGCGGGCGGAGCTGGTG 513
 Db 4 ATGGCGTGTGTGACGAGCTGGCGCTGGGCTGGGACTTCATCAGCGGGCGGAGCTGGTG 63

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 15:05:53 ; Search time 2025 Seconds

(without alignments)
15302.818 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atgagcgtggcgccagca.....ccatcgagatctgcgtctga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthm:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_estc:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vri:*

28: gb_gssl:*

29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	671.2	52.6	804	13	BU053381 UI-M-FCO-
2	630.6	49.5	888	10	BG683070 602651746
3	620.2	48.6	762	13	BU611087 UI-M-FCO-
4	619.6	48.6	746	12	BM964264 UI-M-EQO-

5	619.2	48.6	762	13	BU611098
6	619.2	48.6	786	13	BQ179876
7	617.2	48.4	742	14	CA324262
8	612.6	48.0	758	12	BM944510
9	609.2	47.8	833	14	CA450675
10	607	47.6	719	13	BQ745423
11	606.8	47.6	746	12	BM944369
12	603.8	47.4	730	13	BU701617
13	603.2	47.3	733	12	BM947152
14	596.8	46.8	706	13	BQ179888
15	594.4	46.6	767	13	BU613351
16	591.8	46.4	706	13	BQ745569
17	591.4	46.4	764	13	BQ180145
18	589.8	46.3	731	14	CA317985
19	587.6	46.1	730	13	BU613504
20	587.4	46.1	693	14	CA316498
21	587.2	46.1	696	12	BM951369
22	585.8	45.9	733	12	BM944577
23	583.4	45.8	748	14	CA317267
24	582.6	45.7	698	12	BM950544
25	582	45.6	686	13	BU708991
26	579.2	45.4	720	14	CB526693
27	578.6	45.4	700	12	BM944032
28	577.8	45.3	732	14	CA319385
29	576	45.2	713	14	CA318039
30	571	44.8	684	13	BQ768987
31	569.6	44.7	673	12	BM950706
32	568.6	44.6	736	14	CA319869
33	567.6	44.5	692	14	CA318185
34	557.4	43.7	687	14	CB527765
35	547.4	42.9	675	12	BM944210
36	545.8	42.8	716	14	CA325404
37	541.6	42.5	655	14	CA316925
38	539.4	42.3	736	14	CA328687
39	537.6	42.2	777	12	BM947305
40	536.6	42.1	698	14	CB249777
41	534.6	41.9	629	13	BU709002
42	531.2	41.7	677	12	BM946930
43	526.4	41.3	723	14	CA315154
44	525.2	41.2	684	13	BU612423
45	523.2	41.0	716	13	BU053232

ALIGNMENTS

RESULT 1

BU053381

LOCUS

DEFINITION

IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EMAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.

BU053381

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002


```
CC CC -!- FUNCTION: CONTROLS CELL CYCLE CHECKPOINT. MAY PARTICIPATE IN  
CCC TRANSDUCTION OF THE DNA DAMAGE AND REPLICATIONAL STRESS SIGNALS.  
CCC INITIATES CDC25 PHOSPHATASE BY PHOSPHORYLATING IT, PREVENTING THE  
CCC ENTRY INTO MITOSIS. MAY HAVE A ROLE IN MEIOSIS AS WELL.  
CCC -!- ENZYME REGULATION: KINASE ACTIVITY IS UPREGULATED BY  
CCCC AUTOPHOSPHORYLATION. RAPIDLY PHOSPHORYLATED IN RESPONSE TO DNA  
CCCC DAMAGE AND TO REPLICATION BLOCK (BY SIMILARITY).  
CCC -!- SUBCELLULAR LOCATION: Nuclear.  
CCC -!- SIMILARITY: BELONGS TO THE SER/THER FAMILY OF PROTEIN KINASES.  
CCC CDS1 SUBFAMILY.  
CCC -!- SIMILARITY: Contains 1 FHA domain.  
CCC -----  
CCC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CCC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CCC the European Bioinformatics Institute. There are no restrictions on its  
CCC use by non-profit institutions as long as its content is in no way  
CCC modified and this statement is not removed. Usage by and for commercial  
CCC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CCC or send an email to license@isb-sib.ch).  
CCC -----  
DR EMBL; AF086905; AAC83694.1; --.  
DR HSSP; Q63450; IA06.  
DR MGD; MG1:J355321; Chek2.  
DR InterPro; IPR000253; FHA.  
DR InterPro; IPR000719; Prot_kinase.  
DR InterPro; IPR002290; Ser_thr_pkinase.  
DR Pfam; PF00498; FHA; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD0000001; Prot_kinase; 1.  
DR SMART; SM00240; FHA; 1.  
DR SMART; SM00220; S_TKC; 1.  
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.  
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.  
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.  
DR PROSITE; PS00006; FHA_DOMAIN; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Cell cycle; Phosphorylation; Nuclear protein.  
KW Domain 117 179  
FT FHAI  
FT DOMAIN 224 490 PROTEIN KINASE.  
FT NP_BIND 230 238 ATP (BY SIMILARITY).  
FT BINDING 253 253 ATP (BY SIMILARITY).  
FT ACT_SITE 351 351  
SQ SEQUENCE 546 AA; 61088 MW; A7949EFB5572CDA CRC64;
```

240	MaFeRtCQkVAiKtIISKRFPALGSSREADTAPSVETETELKKL-NHPcIIKIKD-VFD	297
120	TEDCVYFAQSYAPAGOLFPIIIPQVGLPBDTVKRCVQQLGLALDFMHGRQLVHRDIKPN	179
298	AEDYI-VLELMEGGELFDRVVGKNLKEATCKLYFYQMLVAVFYQHLHENGIIHRDLKPN	356
180	VLLDFEE--CRRVKLADFQWTRVVG--CRVKRYSGTIPYTPAPVCQAGRADGLAVDTGVD	235
357	VLLSQSEDC-LKIKTFQSQKLGKTSLMRTLCGTPTYLAPELVNSNGTAGYS--RAVD	413
236	WNAFGLVILFCVLTNFPWEAASGADAFBEFVWQGRILFGLFSQWRRTTEPALRMFQRL	295
414	CNSLGVILFICLSGYPPFSEHKTKVQSLKQDIT--SKYNFIPEWVTVDVSEALDLVKKL	470
296	LALPEPRGPAKEVFR--FLKHELTSF-----LRRRPSHRARKPPGDRPPAAGPLLEA	347
471	LVVDYKARLUTTEALNHPWLQDBYMKKKFODLLVQEKNSVTLPVAPAQTSQKRPLELV	530
348	PG 349	

```

RN  SEQUENCE FROM N.A.
RP  TISSUE=Brain;
RX  MEDLINE=98290545; PubMed=9628591;
RA  Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA  Nomura N., Ohara C.;
RT  "Prediction of the coding sequences of unidentified human genes. IX.
RT  The complete sequences of 100 new cDNA clones from brain which can
RT  code for large proteins in vitro.";
RL  DNA RES. 5:31-39(1998).
CC  -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AB011109; BAA25463.1; -.
DR  HSPF; E24941; IBUH.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_kinase.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00069; pkinase_1.
DR  PRINTS; PR00109; TYRKINASE.
DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKc; 1.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR  PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW  Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW  ATP-binding.
FT  DOMAIN 55 306 PROTEIN KINASE.
FT  NP_BIND 61 69 ATP (BY SIMILARITY).
FT  BINDING 84 84 ATP (BY SIMILARITY).
FT  ACT_SITE 178 178 BY SIMILARITY.
SQ  SEQUENCE 661 AA; 74304 MW; 806F37D52CA4710F CRC64;

Query Match 14.1%; Score 316.5; DB 1; Length 661;
Best Local Similarity 27.5%; Pred. No. 1.3e-12;
Matches 110; Conservative 65; Mismatches 156; Indels 69; Gaps 17;

QY  20 GTAPGAGVLLTDMQA-----LILRTLAADSVTKH-----YELVRLGKGTGKV 67
Db  10 GDRPDLGLGAPSPREAVAGATAALEPRKPGVKRHHKHLKHRYELQETLGRKGTGKV 69
QY  68 DLVYKGTGTQWALKFVNKSTK---LKNFLREVSTINSLSPPFIKVPDVFETEDC 123
Db  70 KQATERFSGRVVAKSTKDKIKDEQDMVHRRRIEMSSJ-NHPHIIISYE-VFENKDK 127
QY  124 YVFAQEVAPAGLDLPIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPENVLLF 183
Db  128 IIVIMEVASKGELYDISERRLSERTRHFFROIIVSAVHYCHNGVVRHDKLENILL- 186
QY  184 DRECRVKLADFGMTR--RVGCRVKRVSGITPYTAPVQVQAGRADGLAVDTGVDVAFGV 241
Db  187 DDNC-NIKIADFGUSNLYQKDKFLQTCGSPLYASPIVNGRPVYRGPEVDS---WALGV 241
QY  242 LIFCVLTGNFWEAASGADAFEFVFWQGRGLPGLPSQWRFTPE--ALRMFORLLALE 299
Db  242 LLYLVVGTWEPD---GFD--HKNLIRQISS-----GEYREPTQSDAGRLIRKMLWN 290
QY  300 PERGPAKEVERF-----LKHELTSERLRP---SHRA-----RKPP 333
Db  291 PDRATTIEDIANHWVNWGYKSSVCCDCLDHDSESPLLARIIDWHHRSTGLQADTEAKM 350
QY  334 GDRPPAGPLRLLEAPGLKRTV-----LTSGSGSRPAPPA 369
Db  351 GLAKPTTSEVMLEQRSLKSKSKENDFAQSQQDAVPESPS 390

```

```

RESULT 13
AAK2_RAT
ID AAK2_RAT STANDARD; PRT; 552 AA.
AC Q09137;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.-)
DE (AMPK alpha-2 chain)
GN PRKA2 OR AMPK2 OR AMPK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94209324; PubMed=7908907;
RA Carling D., Aguan K., Woods A., Verhoeven A.J.M., Beri R.K.,
RA Brennan C.H., Sidebottom C., Davison M.D., Scott J.;
RT "Mammalian AMP-activated protein kinase is homologous to yeast and
RT plant protein kinases involved in the regulation of carbon
RT metabolism.";
RL J. Biol. Chem. 269:11442-11448(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95234757; PubMed=7718624;
RA Gao G., Widmer J., Stapleton D., Teh T., Cox T., Kemp B.E.,
RA Witters L.A.;
RT "Catalytic subunits of the porcine and rat 5'-AMP-activated protein
RT kinase are members of the SNF1 protein kinase family.";
RL Biochim. Biophys. Acta 1266:73-82(1995).
CC -!- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.
CC -!- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q09137-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q09137-2; Sequence=VSP_004949, VSP_004950;
CC Note=Lacks the sequence parts essential for kinase activity and
CC is therefore inactive;
CC -!- TISSUE SPECIFICITY: SKELETAL MUSCLE, LOWER LEVELS IN LIVER, HEART,
CC AND KIDNEY.
CC -!- INDUCTION: BY AMP.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. SNF1
CC SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; Z29486; CAA82620.1; -.
DR  EMBL; U12149; AAA85033.1; -.
DR  PIR; A53621; A53621.
DR  HSP; Q63450; 1A06.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_pkinase.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00069; pkinase; 1.

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [7]
 RN SEQUENCE OF 323-775 FROM N.A. (ISOFORM 7).
 RP TISSUE=Urinary bladder;
 RC Reynolds C.H., Patel U.A., Anderton B.H.;
 RA "Homo sapiens mRNA partial sequence for a protein kinase, STK10,
 RT similar to p78/C-TAK1.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in the specific phosphorylation of microtubule-
 CC associated proteins for tau, MAP2 and MAP4. Phosphorylated Cdc25C
 CC on Ser-216
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=7;
 CC Name=1;
 CC IsoId=P27448-1; Sequence=Displayed;
 CC Name=2; Synonyms=CTAK75a;
 CC IsoId=P27448-2; Sequence=VSP_004944;
 CC Name=3;
 CC IsoId=P27448-3; Sequence=VSP_004942, VSP_004944;
 CC Name=4;
 CC IsoId=P27448-4; Sequence=VSP_004942, VSP_004945;
 CC Name=5;
 CC IsoId=P27448-5; Sequence=VSP_004942;
 CC Name=6; Synonyms=p58;
 CC IsoId=P27448-6; Sequence=VSP_004942, VSP_004943, VSP_004944;
 CC Name=7;
 CC IsoId=P27448-7; Sequence=VSP_004943;
 CC -!- TISSUE SPECIFICITY: Ubiquitous
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
 CC SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
 CC -!- SIMILARITY: Contains 1 UBA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U64205; AAC15093.1; --
 CC EMBL; AF159295; AAD48007.1; --
 CC EMBL; AF387637; AAX82367.1; --
 CC EMBL; AF465413; AAL69982.1; --
 CC EMBL; M80359; AAA59991.1; --
 CC EMBL; BC024773; AAH24773.1; --
 CC EMBL; AF170723; AAD51631.1; --
 CC HSSP; Q63450; IAO6.
 CC Genew; HGNC:6897; MARK3.
 CC MIM; 602678; --
 CC InterPro; IPR001772; Kinase Cterm.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR000449; UBA_domain.
 CC Pfam; PF02149; KAI; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00627; UBA; 1.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC SMART; SM00165; UBA; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50030; UBA; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Alternative splicing.
 FT DOMAIN 56 330 PROTEIN KINASE.
 FT DOMAIN 326 365 UBA.
 FT DOMAIN 727 776 KAI.
 FT NP_BIND 62 70 ATP (BY SIMILARITY).
 FT BINDING 85 85 ATP (BY SIMILARITY).
 FT ACT_SITE 201 201 BY SIMILARITY.
 FT VARSPPLIC 162 184 Missing (in isoform 3, isoform 4, isoform
 FT 5 and isoform 6).
 FT /FTId=VSP_004942.
 FT Missing (in isoform 6 and isoform 7).
 FT /FTId=VSP_004943.
 FT Missing (in isoform 2, isoform 3 and
 FT isoform 6).
 FT /FTId=VSP_004944.
 FT Missing (in isoform 4).
 FT /FTId=VSP_004945.
 FT E -> Q (IN REF. 5).
 FT E -> K (IN REF. 4 AND 5).
 FT R -> K (IN REF. 2).
 FT A -> G (IN REF. 5).
 FT S -> G (IN REF. 1 AND 2).
 FT S -> T (IN REF. 5).
 FT A -> D (IN REF. 5).
 FT N -> T (IN REF. 5).
 FT E -> K (IN REF. 5).
 SQ SEQUENCE 776 AA; 86944 MW; A245496849070098 CRC64;
 Query Match 14.2%; Score 317; DB 1; Length 776;
 Best Local Similarity 28.7%; Pred. NC. 1.4e-12;
 Matches 91; Conservative 64; Mismatches 104; Indels 58; Gaps 13;
 QY 37 QALTLAT-----LAASDVTRK--YEVLRELKGTGKVDLVVYKGTGTMALK 82
 Db 26 QEVTSKTSRGARCRNSIASCADEQPHIGNYRLKTIKGNFAKVLARHLTGREVALK 85
 QY 83 FVNKSK---TKLKNFLREVSITNSLSSPFIKVFDFWFEFETEDCYVFAQEPAGDLFDI 139
 Db 86 IIDTQLNPTSLQKLFREVRIMKIL-NFENIVKLFVEI-ETEKTLYLIMEYASGGEVFDY 143
 QY 140 IPPQVGLPDTVRCVQQ-----LGL-----ALDFHGRQLVHRDIK 176
 Db 144 LVAHGRMKEARSKFRQCGQAGQTIKQVPSFDLLSLMFTFIVSAVQYCHQKRIHRDLK 203
 QY 177 PENVLFPDRECRVKLADFGMTR--VGRVRKVSQGTIPVTAPEVCQAGRADGLAVDTGV 234
 Db 204 AENLLI-DAD-MNIKIADFGFSNEFTVGGKIDTFCSPPYAAPELFGKKYDG----PEV 257
 QY 235 DVWAFGLVIFCVLTGNFPWEAASGADAFPEEFVRQGRGLPGLPSQWRRTPEPALRMFOR 294
 Db 258 DVWSLGVILTVSGSLPFDGQNLKE--LRERVLGRKYRIPFYMS-----TDCENLLKR 309
 QY 295 LLALEPERRGPAKEVFR 311
 Db 310 FLVINPIKRGTLQIMK 326
 RESULT 12
 Y537_HUMAN
 ID Y537_HUMAN STANDARD; PRT; 661 AA.
 AC O60285;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable serine/threonine-protein kinase KIAA0537 (EC 2.7.1.1-).
 GN KIAA0537.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Wältjens I., vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrall B.G., Nurse P.;
 RA "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; AL031543; CAA20833.1; -
 DR PIR; T41587; T41587.
 DR HSP; Q63450; IAO6.
 DR GeneDB SPombe; SPCC74.03c; -
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00030; UBA; 1.
 DR Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 34 285 PROTEIN KINASE.
 FT DOMAIN 304 345 UBA.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 FT ACT_SITE 576 AA; 65996 MW; ES857E8F171E7B50 CRC64;
 SQ SEQUENCE 576 AA; 65996 MW; ES857E8F171E7B50 CRC64;

Query Match 14.2%; Score 317.5; DB 1; Length 576;
 Best Local Similarity 30.8%; Pred. No. 9.5e-13;
 Matches 98; Conservative 56; Mismatches 101; Indels 63; Gaps 15;
 42 RTLAASDVTHY----ELVRE-LGKGTGKVDLVVYKGTGTQKALKFNKSKTKLKNF--- 94
 19 RVLPEPAISRKHIGFYIIRETLGSGSGKVKVLAHYTKQKVALKFKISRLKKSDMHMR 78
 95 -LREVSITNSLSSPFIKVDVNVFEEDCVFAQVAPAGDLFDLIPPQVGLPEDIIVKR 153
 79 VEREISVL-KLLRHPIHIIKLYDVITPTTD-LVMVIEA-GELPFIYVEKRWMTDEGRR 135
 154 CVOQLGLALDFPMHGRQLVHRDIKENVILLDFECRRVKLADFGMTRRV--GCRVKRVSGT 211
 136 FFQIICAEIYCHRRKIVHRDLKENLLDDN--LNVKIADFGLSNMTDGNFKTSKGS 193
 212 IPTAPEVCQGRADGLAVDTGVVDVAFVGLFICVLGTGNFPWEAASGADAFFEEFVRQR 271
 194 PNYAAPEVING----KLYAGPEVDVMSGCVILVYMLVGRLP-----FDDEFI---- 236

QY 272 GRLEGPSPQWR-----FTEP-ALRMFORLLALEPERRGPAKEVFRFLKHELT-S 319
 DB 237 -----PNLFKKVNSCVVVPDPLSPGASQLIRRMIVADPMQR-----ITIQ 277
 QY 320 ELRRRPSHRARKPPGDRP 337
 DB 278 EIRRDPMFNVNLDYLRP 295
 RESULT 11
 MKK3 HUMAN
 ID MKK3_HUMAN STANDARD; PRT; 776 AA.
 AC P27448; O60219; Q8TB41; Q8WK83; Q96RG1; Q9UMY9; Q9UN34;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE MAP/microtubule affinity-regulating kinase 3 (EC 2.7.1.27) (Cdc25C-
 DE associated protein kinase 1) (CTAK1) (C-TAK1) (Serine/threonine
 DE protein kinase p78) (Ser/Thr protein kinase PAR-1) (protein kinase
 DE STRK10).
 DE MARK3 OR CTAK1.
 GN Homo sapiens (Human).
 OS
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=98202387; PubMed=9543386;
 RA Peng C.Y., Graves P.R., Ogg S., Thoma R.S., Byrnes M.J. III, Wu Z.,
 RA Stephenson M.T., Plwnica-Worms H.;
 RA "C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and
 RT promotes 14-3-3 protein binding.";
 RT Cell Growth Differ. 9:197-208(1998).
 RL [2]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP Waggoner S.N., Stephen R., Farrar W.L., Howard O.M.Z.;
 RA "Human serine/threonine protein kinase cTAK1/Kp78/Mark3:
 RT Identification of a novel splice variant and a larger 5'UTR.";
 RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM 4).
 RX MEDLINE=21326471; PubMed=1433294;
 RA Sun T.Q., Lu B., Feng J.J., Reinhard C., Jan Y.N., Fantl W.J.,
 RA Williams L.T.;
 RA "PAR-1 is a Dishevelled-associated kinase and a positive regulator of
 RT Wnt signalling.";
 RT Nat. Cell Biol. 3:628-636(2001).
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORM 5).
 RP Drewes G.;
 RA "Characterization of an alternatively spliced form of MARK3 from human
 RT brain.";
 RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A. (ISOFORM 6).
 RP Maheshwari K.K., Som S., Parsa I.;
 RA Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A. (ISOFORM 3).
 RP TISSUE=Cervix;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_Kinase; 1.
DR SMART: SM00220; S_TK; 1.
DR DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
DR KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
DR DOMAIN 16 29 POLY-HIS.
DR FT DOMAIN 53 305 PROTEIN KINASE.
DR FT NP BIND 59 67 ATP (BY SIMILARITY).
DR FT BINDING 82 82 ATP (BY SIMILARITY).
DR FT ACT SITE 175 175 BY SIMILARITY.
DR FT MOD_RES 208 208 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
DR FT CONFLICT 7 8 PQ -> AR (IN REF. 2).
DR FT CONFLICT 228 234 AGEVDV -> SSVQMKI (IN REF. 2).
DR FT CONFLICT 242 255 YVMGLGRFPDDEF -> GMSCCVDVHSMTSS (IN REF. 2).
DR FT FT 387 387 A -> R (IN REF. 2).
DR FT CONFLICT 416 416 R -> A (IN REF. 2).
DR FT CONFLICT 494 494 S -> L (IN REF. 2).
DR SQ SEQUENCE 620 AA; 70005 MW; 1806C652B5061D2B CRC64;
Query Match 14.5%; Score 324; DB 1; Length 620;

```

QY	53	YELVRELGGTGYGVKDLVVYKGTTGKTWALKFVNKSXTKLKNFL-----REVSTNSLSSS	107
Dd	53	YQLTKLTGEGSGFKVLAKHLGTGQGVALKIINR-KTLAKSDMQGRVEREISYLRLLRH	110
QY	108	PRIIKVFVDVFEETDCVFAQAQEVAPAGDLDFTIPPOVGLPEDTVKRCVQQGLALDFMHG	167
Dd	111	PHIUKYDVI-KSKDEIIIVIEFA-GKELFDYIVQRGMPEDEARFFQOIIAAVEYCHR	168
QY	168	ROLVHRDIKPENLVLLFDRECRRYVKLADFGMTRRV--GCRVRYSGITPY-TAPEVCOAGR	224
Dd	169	HKTIHVRLDKPENLLDDQ--LNVIKIDAFGLSNIMTDGNFLKTSGSPNYNFAPEVI-SGK	225
QY	225	ADGLAVDTGDVWFAGVILFCVLVTGNFPWEAASGADAFFEEFYRWQRGLPGLPSOWRR-	283
Dd	226	LYAGPEVDVWSAGVILIYMLCGRLP-----FDDEFI-----PALFKKI	263
QY	284	FTPE-----ALRMFORLLALEPERRGPAKEV	309
Dd	264	SNGVYTLENYLSAGAKHLLTMLVNPENRTITHEI	299

RESULT 10
 SNFL_SCHPO STANDARD; PRT; 576 AA.
 ID ID SNFL_SCHPO AC 074536;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE SNFL-like protein kinase (EC 2.7.1.-).
 GN SPCC74.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidaoglou J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```

Oy 287 PALRMFORLLALEPERRGPAKEVFR-----FLKHELFSELRRRP----- 325
Db 276 DCENLLKFLILNPSKRGTELEQIMKDRWMNVGHEDDELPKVPVLETTGPRDRDVGNGVL 335
Oy 326 -----SHARKPPGP-----RP-PAAGPLRLAPGP---LKR 353
Db 336 HTEBIOQSLVQORVNEWVATVLLGYKSSEPGDITILKPRPSADLTNNSAPSPSHKVQR 395
Oy 354 TVLTSGSGSRPAPPVAVGSPV 374
Db 396 SV-SANPKQRSSDQAVPAIP 415

RESULT 8
CHK1 HUMAN
ID -CHK1 HUMAN STANDARD; PRT; 476 AA.
AC O14757;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE Serine/threonine-protein kinase Chk1 (EC 2.7.1.-).
GN CHK1 OR CHK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426625; PubMed=9278511;
RA Sanchez X., Wong C., Thoma R.S., Richman R., Wu Z., Pivnicka-Worms H.,
RA Elledge S.J.;
RT "Conservation of the chk1 checkpoint pathway in mammals: linkage of
RT DNA damage to Cdk regulation through Cdc25."
RL Science 277:1497-1501(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98044285; PubMed=9382850;
RA Flagg G., Plug A.W., Dunks K.M., Mundt K.E., Ford J.C.,
RA Quigg M.R.E., Taylor E.M., Westphal C.H., Ashley T., Hoekstra M.F.,
RA Carr A.M.;
RT "Atm-dependent interactions of a mammalian chk1 homolog with meiotic
RT chromosomes."
RL Curr. Biol. 7:977-986(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20184068; PubMed=10717241;
RA Semba S., Ouyang H., Han S.-Y., Kato Y., Horii A.;
RT "Analysis of the candidate target genes for mutation in microsatellite
RT instability-positive cancers of the colorectum, stomach, and
RT endometrium."
RL Int. J. Oncol. 16:731-737(2000).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeng H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC [-] FUNCTION: INVOLVED IN CELL CYCLE ARREST WHEN DNA DAMAGE HAS
CC OCCURRED OR WHEN UNLIGATED DNA IS PRESENT. BINDS TO AND
CC PHOSPHORYLATES CDC25A, CDC25B AND CDC25C. PHOSPHORYLATION OF
CC CDC25C CREATES A BINDING SITE FOR 14-3-3 PROTEIN WHICH INHIBITS
CC CDC25C. THIS PREVENTS ACTIVATION OF THE CDC2-CYCLIN B COMPLEX AND
CC PREVENTS MITOTIC ENTRY. PHOSPHORYLATES CDC25C ON SERINE-216.
CC [-] SUBCELLULAR LOCATION: Nuclear.
CC [-] TISSUE SPECIFICITY: EXPRESSED UBIQUITOUSLY WITH THE MOST ABUNDANT
CC EXPRESSION IN THYMUS, TESTIS, SMALL INTESTINE, AND COLON.
CC [-] PTM: PHOSPHORYLATED.
CC [-] SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AF016582; AAC51736.1; -
CC EMBL; AF032874; AAB88852.1; -
CC EMBL; AB032387; BAA84577.1; -
CC EMBL; AB032377; BAA84577.1; JOINED.
CC EMBL; AB032378; BAA84577.1; JOINED.
CC EMBL; AB032379; BAA84577.1; JOINED.
CC EMBL; AB032380; BAA84577.1; JOINED.
CC EMBL; AB032381; BAA84577.1; JOINED.
CC EMBL; AB032382; BAA84577.1; JOINED.
CC EMBL; AB032383; BAA84577.1; JOINED.
CC EMBL; AB032384; BAA84577.1; JOINED.
CC EMBL; AB032385; BAA84577.1; JOINED.
CC EMBL; AB032386; BAA84577.1; JOINED.
CC EMBL; AF527555; AAM78553.1; -
CC EMBL; BC004202; AAH04202.1; -
CC EMBL; BC017575; AAH17575.1; -
CC PDB; 1IA8; 18-APR-01.
CC Genew; HGNC:1925; CHEK1.
CC GK; O14757; -
CC MIM; 603078; -
CC GO; GO:0005711; C:meiotic chromosome; TAS.
CC GO; GO:0008181; F:tumor suppressor; TAS.
CC GO; GO:0000077; P:DNA damage checkpoint; TAS.
CC GO; GO:0000776; P:gametogenesis; TAS.
CC GO; GO:0007276; P:meiotic recombination; TAS.
CC GO; GO:0007131; P:negative regulation of cell proliferation; TAS.
CC GO; GO:0008285; P:oncogenesis; TAS.
CC GO; GO:0007048; P:regulation of CDK activity; TAS.
CC GO; GO:0000079; P:regulation of CDK activity; TAS.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC Pfam; PF00069; pkinase; I.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
CC DNA damage; Nuclear protein; Phosphorylation; 3D-structure.
CC DOMAIN 9 265
CC NP BIND 15 23
CC FT BINDING 38 38
CC FT ACT SITE 130 130
CC BY SIMILARITY.
CC SEQUENCE 476 AA; 54419 MW; 0ABD0FAB67F49F67 CRC64;

```


the checkpoint response in *S. pombe*.";
Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.

RT STRAIN=972;
RX MEDLINE=21848401; PubMed=1859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Hollroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy S., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.U., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of *Schizosaccharomyces pombe*.";
Nature 415:871-880(2002).

-1- FUNCTION: HAS A ROLE IN THE DNA REPLICATION-MONITORING S/G2
CHECKPOINT SYSTEM. IT IS RESPONSIBLE FOR BLOCKING MITOSIS IN THE
S PHASE. IT MONITORS DNA SYNTHESIS BY INTERACTING WITH DNA
POLYMERASE ALPHA AND SENDS A SIGNAL TO BLOCK THE ONSET OF MITOSIS
WHILE DNA SYNTHESIS IS IN PROGRESS.
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CDS1
SUBFAMILY.
-1- SIMILARITY: Contains 1 FHA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X85040; CAA59410.1; -;
DR EMBL; AJ222869; CAA11019.1; -;
DR EMBL; AL109736; CAB52158.1; -;
DR FRR; S58882; S58882;
DR HSSP; Q00534; IBI7.
DR GeneDB_SFombe; SPCC18B5.11c; -;
DR InterPro; IPR000253; FHA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00006; FHA_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
Cell cycle.
KW DOMAIN 60 116 FHA.
FT DOMAIN 167 433 PROTEIN KINASE.

FT NP_BIND 173 181 ATP (BY SIMILARITY).
FT BINDING 196 196 ATP (BY SIMILARITY).
FT ACT_SITE 294 294 R -> G (IN REF. 1).
FT CONFLICT 61 61 F -> C (IN REF. 1).
FT CONFLICT 238 238 F -> C (IN REF. 1).
SQ SEQUENCE 460 AA; 52014 MW; 4CEB963D3376DB54 CRC64;

Query Match 15.0%; Score 335; DB 1; Length 460;
Best Local Similarity 30.5%; Pred. No. 6.2e-14;
Matches 90; Conservative 56; Mismatches 121; Indels 28; Gaps 11;

QY 52 HYELVRELKGTGKGVLDLVVYKGTGKMAKLVNKSATKLN-----FLREVSTNS 103
Db 166 HYIIRTLGSGTFAVVKLAVERNKGWYAIKINKRILLTSSEKRATEMFOREIDILKS 225

QY 104 LSSSPFIKVFVVFETDCYFAQAYAPAGDLFDIIPQVGLPDPVTRKCVQQLGLALD 163
Db 226 L-HHPGVVQCHE-IFENDDLEFVMEYVEGDLMDFLIANGSIDEQDCKPLKQLLETL 283

QY 164 FMHGROLVHRDIKPNVLLDRECRVRVLADPGMTRV---GCRVKVSGTIPYTAPEVC 220
Db 284 HLHKQGVTHRDIKPENILITND--FHLKISDFGLAKVIHGTGTFTFCCTMGYLAPEVL 341

QY 221 QAGRA--DGLAVDTGVYDWFAGVLIPLVTGTFNFEWAASGADAFEFVFWQGRPLGLP 278
Db 342 KSKNVLDG-GYDDKVDIWSLGCVLVMTLSTPFASSSOAKCI--ELI--SKGAYPIEP 396

QY 279 SQWRRETEPALRNFORLALALEPGRGPAKEVFRFLKHELTSELRPRPSHRARPPP 333
Db 397 LLENBIESEGIDILNMLNPEKRISSEA---LQHPWFYTV---STHEHRTTP 445

RESULT 6

CHK1 MOUSE
ID CHK1 MOUSE STANDARD; PRT; 476 AA.
AC Q35280; O54925;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine-protein kinase Chk1 (EC 2.7.1.-).
GN CHEK1 OR CHK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=97426625; PubMed=9278511;
RA Sanchez Y., Wong C., Thoma R.S., Richman R., Wu Z., Pivnicka-Worms H.,
RA Elledge S.J.;
RT "Conservation of the Chk1 checkpoint pathway in mammals: linkage of
DNA damage to Cdk regulation through Cdc25.";
RL Science 277:1497-1501(1997).
RN [2]
RP SEQUENCE OF 204-450 FROM N.A.
RX MEDLINE=98044285; PubMed=9382850;
RA Flagg G., Plug A.W., Dunks K.M., Mundt K.E., Ford J.C.,
RA Quiggle M.R.E., Taylor E.M., Westphal C.H., Ashley T., Hoekstra M.F.,
RA Carr A.M.;
RT "Atm-dependent interactions of a mammalian chk1 homolog with meiotic
chromosomes.";
RL Curr. Biol. 7:977-986(1997).
CC -1- FUNCTION: INVOLVED IN CELL CYCLE ARREST WHEN DNA DAMAGE HAS
OCCURRED OR WHEN UNLIGATED DNA IS PRESENT. BINDS TO AND
PHOSPHORYLATES CDC25A, CDC25B AND CDC25C. PHOSPHORYLATION OF
CDC25C CREATES A BINDING SITE FOR 14-3-3 PROTEIN WHICH INHIBITS
CDC25C. THIS PREVENTS ACTIVATION OF THE CDC2-CYCLIN B COMPLEX AND
PREVENTS MITOTIC ENTRY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: FOUND IN ALL ADULT TISSUES TESTED. ELEVATED
EXPRESSION IN TESTIS, LUNG AND SPLEEN. 15.5 DAY OLD EMBRYOS SHOW
UBIQUITOUS EXPRESSION WITH STRONG EXPRESSION IN BRAIN, LIVER,
CC KIDNEY, PANCREAS, INTESTINE, THYMUS AND LUNG.

Subsequence (DEC-1996) to the EMBL/GenBank/DBJ databases.

```

RL  [3]
RN  SEQUENCE FROM N.A.
RP  STRAIN=cv. Columbia;
RX  MEDLINE=21016720; PubMed=11130713;
RA  Salanoubat M., Lemcke K., Rieger M., Anseorge W., Unseld M.,
RA  Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA  Delisney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA  De Simone V., Choisine N., Artiguenave F., Robert C., Brottier P.,
RA  Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA  Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA  Wurmbach E., Dronek H., Efrle H., Jordan N., Brandt P., Nyakatura G.,
RA  Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Simonati B.,
RA  Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA  Conrad A., Hornschke K., Kauer G., Boegner T.-H., Nordaiek G.,
RA  Reichelt J., Schaefer M., Schoen O., Barques M., Terol J., Climent J.,
RA  Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA  Cooke R., Laudie M., Berger-Lilauro C., Furnelle B., Masuy D.,
RA  de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA  Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA  Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA  Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA  Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA  Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA  Pai G., Millischer J., Sellers P., Gill J.E., Feldblyum T.V.,
RA  Preuss D., Lin X., Niemman W.C., Salzberg S.L., White O., Venter J.C.,
RA  Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA  Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA  Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA  Nakayama S., Nakazaki N., Shimpoto M., Takeuchi C., Wada T.,
RA  Watanabe A., Yanada M., Yasuda M., Tabata S.,
RA  "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT  thaliana."
RL  Nature 408:820-822 (2000).
RN  [4]
RN  SEQUENCE OF 1-19 FROM N.A.
RP  STRAIN=cv. Columbia;
RX  MEDLINE=95115691; PubMed=7816049;
RA  le Guen L., Thomas M., Kreis M.;
RA  "Gene density and organization in a small region of the Arabidopsis
RT  thaliana genome."
RL  Mol. Gen. Genet. 245:390-396(1994).
CC  -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION
CC  CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC  HIGHER PLANTS.
CC  -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, SHOOTS AND LEAVES.
CC  -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  SNF1 SUBFAMILY.
CC  -!- SIMILARITY: Contains 1 UBA domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M93023; AAA32736.1; -
CC  EMBL; X79707; CAA56146.1; -
CC  EMBL; AC008261; AAF26165.1; -
CC  EMBL; X94757; CAA64384.1; -
CC  FIR; JCI4446; JCI446.
CC  HSP; Q63450; 1A06.
CC  InterPro; IPR001772; Kinase Cterm.
CC  InterPro; IPR000719; Prot_kinase.
CC  InterPro; IPR002290; Ser_Thr_kinase.
CC  InterPro; IPR001245; Tyr_kinase.
CC  InterPro; IPR000449; UBA_domain.
CC  Pfam; PF02149; KAI; 1.
CC  Pfam; PF00069; pkinase; 1.
CC  Pfam; PF00627; UBA; 1.
CC  PRINTS; PR00109; TYRKINASE.

```

```

DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  SMART; SM00165; UBA; 1.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR  PROSITE; PS00031; PROTEIN KINASE DOM; 1.
DR  PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR  PROSITE; PS00030; UBA; 1.
KW  Transferase; Serine/threonine-protein kinase; ATP-binding;
KW  Phosphorylation.
FT  DOMAIN 19 271 PROTEIN KINASE.
FT  NP_BIND 292 332 UBA. (BY SIMILARITY).
FT  BINDING 25 33 ATP (BY SIMILARITY).
FT  ACT_SITE 48 48 ATP (BY SIMILARITY).
FT  ACT_SITE 142 142 BY SIMILARITY.
FT  MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ  SEQUENCE 512 AA; 58173 MW; 5A18655A0AA506DF CRC64;

Query Match 15.0%; Score 336.5; DB 1; Length 512;
Best Local Similarity 30.3%; Pred.No. 5.6e-14;
Matches 97; Conservative 59; Mismatches 111; Indels 53; Gaps 13;

QY 20 GTAPGPGAGVPLLTEDMQALTLTLAASDVTKHELVRELKGTGKVDLVVYKGTGTM 79
DB 3 GSTGGRGV-----ESLPYKLGRTLGISFGKVKAEHALTGKV 45
QY 80 ALKFVNKSTK----LKNFLREYSITNSLSSPFIKVFVDVFETEDCYVFAQYAPAGD 135
DB 46 AIKILNRKIKNMEMBEKVRREIKILR-LFMHPHILRLVEVETPTDIYL-VMEYVNSGE 103
QY 136 LFLIIPQVGLPDTVRCVQOGLALDFMHGQLVHRDIKPNVLLFDECRVRKLAADF 195
DB 104 LFDYIVKEKGLQDEARNFFQIIISGVEYCHRNWVHRDLKPNLL-DSKC-NVKIADF 161
QY 196 GMTR--RVGCRVRSVTIETAPETVCQAGRADGLAVDTGVDMVAFGLVFCVLTGNFPW 253
DB 162 GLSNIMEDGHFLXTCGSPNYAPEVI-SK---LYAGPEVDVWSCGVIIYALCGTLFF 217
QY 254 EAAGADAFEEFVRVWRGRLPGLPQWRRTFPALR-MFORLLALEPERGPAKEVFRF 312
DB 218 D-----DENIPNLFKKIGGIYLPKH---LSFGARDLIPRLMVDP----- 256
QY 313 LKHELTSELRRRSHRARP 332
DB 257 MKRVTPETIRQHPWFOAHP 276

RESULT 5
CDL1_SCHPO STANDARD; PRT; 460 AA.
AC Q09170; O42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein kinase cds1 (EC 2.7.1.-) (Checkpoint kinase cds1).
GN CDS1 OR SPC1185.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95240713; PubMed=7723827;
RA Murakami H., Okayama H.;
RT "A kinase from fission yeast responsible for blocking mitosis in S
RT phase."
RL Nature 374:817-819 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA Lindsay H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
RA Murray J.M., Osman F., Walworth N., Carr A.M.;
RT "S-phase specific activation of Cds1 kinase defines a sub-pathway of

```

```

FT PPEALMAALROATAACRCROPOPIIFLACILHGGAGGPEPL
FT SHEVEVCQLRPGLRGVLPFRVAGTALAFRTILVIRISNDL
FT EL -> TLDPRKQRCVSGSLPQGSKIRSQINURESG
FT DLRSQVAIYLGIKRRPPGCSDSPEV (in isoform
FT 2).
FT /FTID=VSP 004946.
FT CONFLICT 70 70
FT F -> S (IN REF. 2).
SQ SEQUENCE 752 AA; 82519 MW; 4B430FFD2B150E7A CRC64;

Query Match 15.3%; Score 341.5; DB 1; Length 752;
Best Local Similarity 26.3%; Pred. No. 4.1e-14;
Matches 109; Conservative 70; Mismatches 155; Indels 81; Gaps 15;

QY 52 HYELVRELKGTGYGVLDVAVYKGTGKMAKLVNKS---TKLKNFREVSTNLSLSSP 108
DB 58 NYRLRLTIGKGNFAKVKLARHILTGREVAIKIITQINPSSLOKLFREVIMKGL-NHP 116
QY 109 FIIKVPDVVFETDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHR 168
DB 117 NIVKLEVI-ETETLYLVMEYASAGVEDYLVSHGRMKKEAKRQIVSAVHYCHOK 175
QY 169 QLVHRDIKPNVLLFDRECRVKLADFGWTRR--VGCRVKRVSGTIPYTAPEVCQAGRA 226
DB 176 NIVHRDLKAENLL--DAEA-NIKIADFGSNEFTLGSKLDTFCGSPPYAADELFGKKYD 233
QY 227 GLAVDTGVDVWAFGLVPCVLTGNFPEAASGADAFEEFVWQGRGLPGLPSQWRPTE 286
DB 234 G---PEVDIWSGLVITYVSGSLPFDGHN---LKELRVLRGKY-RVFFYMSYDCE 284
QY 287 PALRMFORLLALEPERRGPAKEVFR----- 311
DB 285 SILR---RFLVNPAXKRTLLQIMKDKWINIGYGEELKPYTEPEDEFGDTKRIEVMVGM 341
QY 312 -----FLKHETLSLRPPS-----HRAKPPGDR-PPAAGPRLRLAPGLPKRTVTUTES 359
DB 342 GYTBREKESUTSQKNEVTATYLLGKTEBEGDGRGAPGLARVAPSDDTTNGTSSSK 401
QY 360 GSGSRPAPPVAGSV-----PLPVVPVVPVVPVVPVVPVPPGLAPQGPGR 402
DB 402 GTSHSGKGRSSSYHRRHSDSCGSPALH-PKRSPTSTGAELEKELPGR 455

RESULT 3
SNF1_CANTR STANDARD; PRT; 619 AA.
AC O94168;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN SNF1.
OS Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanai T., Ogawa K., Ueda M., Tanaka A.;
RT "Genetic evaluation of the function of SNF1 in Candida tropicalis.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
CC PROTEIN SNF4. COULD PHOSPHORYLATES CATB (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB024535; BAA75889.1; -.
DR HSSP; Q63450; IAC6.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Chr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
FT DOMAIN 17 29 POLY-HIS.
FT DOMAIN 52 303 PROTEIN KINASE.
FT NP_BIND 58 66 ATP (BY SIMILARITY).
FT BINDING 81 81 ATP (BY SIMILARITY).
FT ACT_SITE 174 174 BY SIMILARITY.
FT MOD_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 619 AA; 70323 MW; 0PCF1FC3DCE706D7 CRC64;

Query Match 15.1%; Score 337.5; DB 1; Length 619;
Best Local Similarity 33.5%; Pred. No. 5.9e-14;
Matches 92; Conservative 59; Mismatches 77; Indels 47; Gaps 13;

QY 53 YELVRELKGTGYGVLDVAVYKGTGKMAKLVNKS---TKLKNFREVSTNLSLSS 107
DB 52 YQIKTLGEGSGFKVLAQHVGVGQKVALKINR-KTLAKSDMQGRVERISYLR-ILRH 109
QY 108 FFIKVPDVVFETDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHR 167
DB 110 PHIKLYDVIT-KSKDEILIMVIEFA-GKELEFDYIVQGRKPEDEARRFFQIIAAVEYCHR 167
QY 168 RQLVHRDIKPNVLLFDRECRVKLADFGWTRR--GCRVKRVSGTIPYTAPEVCQAGRA 225
DB 168 HKVHRDLKPEENLLDQ--LNVKIADFGUSNTDGNFLKTCGSPNYAAPEVI-SGK- 223
QY 226 DGLAVDTGVDVWAFGLVPCVLTGNFPEAASGADAFEEFVWQGRGLPGLPSQWR-- 283
DB 224 --LYAGEVDVWSSGVLYVWLCGRLP-----FDDEFI-----PALFKKIS 262
QY 284 ---FTEP-----ALRMFORLLALEPERRGPAKEV 309
DB 263 NGVYTLFNLSPGAKHLLTLMVNVNPLRTITHEI 297

RESULT 4
KIL0_ARATH STANDARD; PRT; 512 AA.
AC Q38997; O04728; Q39076;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SRP-2003 (Rel. 42, Last annotation update)
DE SNF1-related protein kinase KIN10 (EC 2.7.1.-) (AKIN10).
GN KIN10 OR SKIN10 OR AT3G01090 OR T4P13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013041; PubMed=1339373;
RA le Guen L., Thomas M., Bianchi M., Halford N.G., Kreis M.;
RT "Structure and expression of a gene from Arabidopsis thaliana
RT encoding a protein related to SNF1 protein kinase.";
RL Gene 120:249-254(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC SPRAIN-cv. Columbia;
RA Lessard P., Kreis M., Thomas M.;

```

```

FT NP_BIND 132 140 ATP (BY SIMILARITY).
FT BINDING 155 155 ATP (BY SIMILARITY).
FT ACT_SITE 251 RPEED -> VAVPY (in isoform b).
FT VARSPLIC 402 406 /FTid=VSP.004898.
FT VARSPLIC 408 469 Missing (in isoform b).
FT VARSPLIC /FTid=VSP.004899.
SQ SEQUENCE 469 AA; 54158 MW; A71616233ABD05A9 CRC64;
Query Match 19.4%; Score 433.5; DB 1; Length 469;
Best Local Similarity 29.3%; Pred. No. 5.1e-20;
Matches 99; Conservative 79; Mismatches 141; Indels 19; Gaps 5;

QY 38 ALTERTLAASDVTHKHYELVRELKGTGKVDLVVYKGTGTMALKFVFNKSKTKLKNFLRE 97
DB 111 SVTMEIKKVDLEBYVTVNKGITGREGFKIAEHKQSKQRIAKTFFRPQTQADFPVE 170
QY 98 VSIINSSSPFIKVDVVPETED--CYVPAQYAPAGDLFDIIPP--QVGLPEDTVKR 153
DB 171 YNYSFFLSPQNIIDTVEGMFQSSDDTAIFYVQFCFRASREAVEATNQAGIGEANTKK 230
QY 154 CVOQLGLALDFHGRQVHRDIKPNVLLFD-RECRVVKLADFGMTERRVGRVGRVSGTI 212
DB 231 VFAAVLSAIFRMDENLVHRNKAENILIFDANDYSKVYKVTDFGLTRKVDITTKYLEVYN 290
QY 213 PYTAPEVCQAGRADGLAVDTGVVWAGVLIFCVLTGNFPEWAAAGADAFEEFVRWORG 272
DB 291 NYHAAELCDTVNNEKLVNKSSTDIWALGIIPFYCMKGFFWQKASIMCKPYWEWEQWLKR 350
QY 273 RLPLGSRMRRTPEPALRMFORLALPERRGPAKEVREL-KHELTSELRRRSHRARK 331
DB 351 KNPALPKKFPFSEKALFKKSTTPFKDKRTWTKMRKLAKEKLKLSVK----- 401
QY 332 PGDRPPAAGPLRLAEPGLKRTVLTESGSGSRPAPPA 369
DB 402 ----RPEEDYVVMIDTASKSRQTATSSGEPQDSAPPA 435

RESULT 2
MRK4 HUMAN
ID MRK4 HUMAN STANDARD; PRT; 752 AA.
AC Q96L34; Q96JG7; Q9BYD8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE MAP/microtubule affinity-regulating kinase 4 [EC 2.7.1.27]
DE DE (MAP/microtubule affinity-regulating kinase like 1).
DE MARK4 OR MARKL1 OR KIAA1860.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RA Drewes G., Mandelkow E.M.;
RC TISSUE=Brain;
RT "MARK4, homologue of MARK1, MARK2 and MARK3.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A. (ISOFORM 2).
RA Kato T., Satoh S., Okabe H., Kitahara O., Ono K., Kihara C.,
RC Tanaka T., Tsunoda T., Yamaoka Y., Nakamura Y., Furukawa Y.;
RT "Isolation of a novel human gene, MARKL1, homologous to MARK3 and its
RL involvement in hepatocellular carcinogenesis.";
RN [3]
Neoplasia 3:4-9(2001).
SEQUENCE FROM N.A. (ISOFORM 1).
RA Beghini A., Magnani I., Roversi G., Piepoli T., Diterlizzi S.,
RC Pollo B., Conti A.M.F., Cowell J.K., Finocchiaro G., Larizza L.;
RT "Neural progenitor-restricted isoform of MARKL1 gene is upregulated by

```

```

RT 19q13 amplification in human glioblastoma.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
RN [5]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Thyroid;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuhara Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96L34-1; Sequence=Displayed;
CC Name=2; Synonyms=MARKL1S;
CC IsoId=Q96L34-2; Sequence=VSP.004946;
CC -|- TISSUE SPECIFICITY: Ubiquitous. Isoform 2 is brain-specific.
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
CC SUBFAMILY.
CC -|- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -|- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AY057448; AAL23683.1; -
EMBL; AB049127; BAB39380.1; -
EMBL; AY120867; AAM55491.1; -
EMBL; AK075272; BAC11510.1; -
EMBL; AB058763; BAB47489.1; ALT_INIT.
HSP; Q63450; IA06.
Genew; HGNC:13538; MARK4.
DR InterPro; IPR001772; Kinase_Cterm.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PSS0030; UBA; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 59 310 PROTEIN_KINASE.
FT DOMAIN 324 368 UBA.
FT DOMAIN 703 752 KAI.
FT NP_BIND 65 73 ATP (BY SIMILARITY).
FT BINDING 88 88 ATP (BY SIMILARITY).
FT ACT_SITE 181 181 BY SIMILARITY.
FT VARSPLIC 628 752 ADEPERIGGEVTSCHLPWDQTETAPRLLPFWNSVLTSSR

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:50:56 ; Search time 17 Seconds
(without alignments)
1172.901 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPFPRLTCCGPG.....GRADSKGQVVLATAIEICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	433.5	19.4	469	1	Q11090 caenorhabdi
2	341.5	15.3	752	1	Q96134 homo sapien
3	337.5	15.1	619	1	Q94168 candida tro
4	336.5	15.0	512	1	Q89997 arabidopsi
5	335	15.0	460	1	Q09170 schizosacch
6	333	14.9	476	1	Q35280 mus musculu
7	330.5	14.8	774	1	Q05512 mus musculu
8	328	14.6	476	1	Q14757 homo sapien
9	324	14.5	620	1	Q52497 candida alb
10	317.5	14.2	576	1	Q74536 schizosacch
11	317	14.2	776	1	P27448 homo sapien
12	316.5	14.1	661	1	Q60285 homo sapien
13	315	14.1	552	1	Q09137 rattus norv
14	314.5	14.0	546	1	Q92265 mus musculu
15	311.5	13.9	1051	1	Q70405 mus musculu
16	311	13.9	502	1	Q02723 scalae cere
17	309.5	13.8	414	1	Q9655 homo sapien
18	309.5	13.8	735	1	Q63531 rattus norv
19	309	13.8	552	1	P54646 homo sapien
20	309	13.8	633	1	P06782 saccharomyc
21	307	13.7	543	1	Q06017 homo sapien
22	306.5	13.7	476	1	Q61267 drosophila
23	306.5	13.7	611	1	Q00372 candida gla
24	304	13.6	735	1	Q5418 homo sapien
25	303.5	13.6	779	1	Q60670 mus musculu
26	303	13.5	1081	1	Q12236 saccharomyc
27	302	13.5	776	1	Q9rlu5 rattus norv
28	302	13.5	1142	1	Q12263 saccharomyc
29	301.5	13.5	786	1	P57059 homo sapien
30	301	13.4	309	1	Q9uqb9 homo sapien
31	301	13.4	733	1	Q9wt13 mus musculu
32	300.5	13.4	740	1	P18112 homo sapien
33	300.5	13.4	740	1	P18654 mus musculu

34 299.5 13.4 397 1 S17A_RABIT
35 299.5 13.4 580 1 SRK1_SCHPO
36 298 13.3 856 1 UN51_CAEEL
37 296 13.2 548 1 AAK1_RAT
38 296 13.2 550 1 AAK1_HUMAN
39 294.5 13.2 1050 1 ULK1_HUMAN
40 294 13.1 733 1 KGA2_HUMAN
41 293.5 13.1 481 1 AKT2_MOUSE
42 293.5 13.1 740 1 DCK1_HUMAN
43 293.5 13.1 756 1 DCK1_MOUSE
44 292.5 13.1 481 1 AKT2_HUMAN
45 291.5 13.0 433 1 DCK1_RAT

ALIGNMENTS

RESULT 1
YMY3_CAEEL STANDARD; PRT; 469 AA.
AC Q11090; Q8MQF5; Q8MQF6;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine/threonine-protein kinase ColC4.3 in chromosome X
DE (EC 2.7.1.1-).
GN ColC4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nelson J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=Q11090-1; Sequence=Displayed;
CC Name=b;
CC IsoId=Q11090-2; Sequence=VSP_004898, VSP_004899;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U41025; AAM51507.1; -;
CC EMBL; U41025; AAM51508.1; -;
CC WormPep; C01C4.3a; CE30845.
CC WormPep; C01C4.3b; CE30846.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00219; Ty_Kc; 1.
CC PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE_ST; FALSE NEG.
CC PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
CC ATP-binding; Alternative splicing.
CC DOMAIN 126 397
CC PROTEIN KINASE.

AAE32027;
24-MAR-2003 (first entry)
Human kinase and phosphatase (KPP)-8.
Human; kinase and phosphatase; KPP; cardiovascular; hypertension; myocardial infarction; angina pectoris; Alzheimer's disease; epilepsy; acquired immune deficiency syndrome; AIDS; Grave's disease; diabetes; neurological; Parkinson's disease; cirrhosis; psoriasis; gene therapy; hypercholesterolemia; anticonvulsant; hepatotropic; lipid myopathy; cell proliferative disorder; cancer; cardiac; neuroprotective; enzyme; nontropic; ophthalmological; anorectic; cytostatic; cataract; obesity.
Homo sapiens.
Key Location/Qualifiers
Peptide 9..39 /label= Signal_peptide
Protein 40..414 /note= "Human mature KPP protein"
Peptide 9..38 /label= Signal_peptide
Protein 39..414 /note= "Human mature KPP protein"
Domain 98..349 /note= "Protein kinase domain"
Domain 271..298 /note= "Transmembrane domain"
WO200283709-A2.
24-OCT-2002.
05-APR-2002; 2002WO-US10818.
06-APR-2001; 2001US-282119P.
13-APR-2001; 2001US-283588P.
13-APR-2001; 2001US-283759P.
20-APR-2001; 2001US-285589P.
27-APR-2001; 2001US-287036P.
27-APR-2001; 2001US-287037P.
04-MAY-2001; 2001US-288608P.
04-MAY-2001; 2001US-288712P.
09-MAY-2001; 2001US-289909P.
17-MAY-2001; 2001US-292246P.
(INCY-) INCYTE GENOMICS INC.
Recipon SA, Burrill JD, Marcus GA, Zingler KA, Tang YT, Thornton M, Borowsky ML, Baughn MR, Burford N, Lee SY, Bandman O; Hafalia AJA, Yao MG, Ramkumar J, Wallia NK, Lu DAM, Arvizu CS; Ison CH, Ding L, Lu Y, Gururajan R, Walsh RT, Gandhi AR; Swarnakar A, Forsythe IV, Yue H, Au-young JK, Elliott VS, Lee S; WPI: 2003-092995/08.
N-PDSB; AAD49449.
New human kinases and phosphatases (KPP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant KPP expression, e.g. hypertension, cancer, AIDS, diabetes, cataract, psoriasis, obesity -
Claim 1; Page 162-163; 195pp; English.
The invention relates to human kinases and phosphatases (KPP), and their corresponding nucleic acids. The polypeptides and polynucleotides of the invention are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of KPP, such as cardiovascular (e.g. hypertension, myocardial infarction, angina pectoris), immune (e.g. acquired immune deficiency syndrome (AIDS), Grave's disease, diabetes), neurological (e.g. Parkinson's

disease, Alzheimer's disease, epilepsy), disorders affecting growth and development (e.g. cirrhosis, psoriasis, cataract), lipid (e.g. hypercholesterolaemia, obesity, lipid myopathies), cell proliferative disorders, or cancer. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of KPP. The KPP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. KPP DNA is used in gene therapy. The present sequence is human KPP protein.

CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX

SQ Sequence 414 AA;

Query Match 22.7%; Score 508.5; DB 24; Length 414;
Best Local Similarity 37.1%; Pred. No. 4.1e-35;
Matches 119; Conservative 38; Mismatches 137; Indels 27; Gaps 4

QY 19 PGTAPGCGAG-----VELLTD-----MQALTTLTAAASDVTKHYE 54
DB 40 PGLPPTQGAGASVAVTNLSRTOKVARVEDTATALQLRLVELTTTSRTPVRSLRDQVH 99
QY 55 LVRELGGTYGVKVDLVVYGVTGWALKFVNKSKTKLNFLREVSITNSLGSSPFIIKV 114
DB 100 LIKKLGSSGRVLIAOPHQGGPAVALKLRLDLVLRSTFDUREFCVGRCSAHPGLLOTL 159
QY 115 DVVFETEDCVFAQEYAPAGDLFIIPFGVLPEDTVKRCVQQGLGALDFMHGROLVHRD 174
DB 160 AGPLQTPRYFAFAQEYAPCGDLSGMQLQR-GLPELLVKRVVAQLAGALDFLHSRGLVHAD 218
QY 175 IKPENVLLEDRECRRVKLADFGMTTRRYGCRVKRYSGTIPTYAPEVCQAGRADGLAVDTGV 234
DB 219 VKEDNVLFVPDCSRVALGDGLTRPGSPTPAPPVELFTAPPELCLLLPPDTLPFRPAV 278
QY 235 DVMAFGYLIFCVLTGNFPWBAAGADAFFEEFVRW--QRGELGPLSQWRFFEAFALRMF 292
DB 279 DSNGLGVLLFCATACFPWDVALAPNEFEFAFAGVTTKPQPQPPPDQCFAPPALALL 338
QY 293 QRLLALEPERRGPAKEVFRFL 313
DB 339 QGLLDLPETRSPPLAVLDLFL 359

RESULT 14
ABB61118
ID ABB61118 standard; Protein; 577 AA.
XX ABB61118;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 10146.
XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX OS Drosophila melanogaster.
XX PN W0200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX ER 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX FI Venter JC, Adams M, Li PWD, Myers EW;
XX

	Query Match	24.4%; Score 546.5; DB 23;	Length 348;
	Best Local Similarity	42.9%;	Pred.No. 1.8e-38;
	Matches 117; Conservative	42; Mismatches 111;	Indels 3; Gaps 3;
QY	34 EDNQALTLRLTAAADVTXHYELVRELGGKYYGVDLVVYGKTGYMALKFVNKSATKLKN	93 : :	

CC of the activity of the kinase. The proteins, nucleic acids
 CC antibodies and modulators are useful for diagnosing, treating or
 CC monitoring cancer or aberrant cellular proliferation and/or
 CC differentiation such as lung, ovarian or brain cancer, pain (e.g.
 CC inflammation or infection) or metabolic disorder (e.g. obesity, anorexia
 CC nervosa, cachexia, lipid disorders and diabetes) or brain disorder
 CC (e.g. Alzheimer's disease, Pick's disease, Parkinsonism, Huntington's
 CC disease), neural tube defects, cerebrovascular diseases (e.g. hypoxia,
 CC ischaemia, infarction), AIDS (acquired immunodeficiency syndrome) -
 CC associated myopathy, multiple sclerosis, disorders of the heart
 CC (e.g. heart failure, angina pectoris, myocardial infarction),
 CC blood vessel disorders (e.g. arteriovenous fistula, vasculitis,
 CC syphilitic aneurysms), disorders of the bone (e.g. osteoporosis,
 CC cirrhosis, hypoparathyroidism, hyperparathyroidism, tropical sprue,
 CC idiopathic hyperglycaemia, hematopoietic disorders such as autoimmune
 CC disease including rheumatoid arthritis, osteoarthritis, multiple
 CC sclerosis, systemic lupus erythematosus, psoriasis, conjunctivitis,
 CC ulcerative colitis, asthma, allergic asthma, autoimmune uveitis,
 CC aplastic anaemia, Grave's disease, graft-versus-host disease; liver
 CC disorders e.g. Gaucher's disease, Wilson's disease; and platelet
 CC disorders. Numerous other diseases that can be treated or prevented
 CC are listed in the specification. The present sequence is the protein
 CC kinase 32374.
 XX
 SQ Sequence 346 AA;

Query Match 81.4%; Score 1822; DB 23; Length 346;
 Best Local Similarity 100.0%; Pred. No. 8.9e-148;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 MALFKFNKSTKLKFNLFREYSITNSLSSTPFIKVFDFWETEDCYVFAQYAPAGDLFD 138
 Db 1 MALFKFNKSTKLKFNLFREYSITNSLSSTPFIKVFDFWETEDCYVFAQYAPAGDLFD 60
 QY 139 IIPQVGLPDTVKRCVQQLGLALDFMHGRLVHRDIKPNVLLFDECRVVKLADFGMT 198
 Db 61 IIPQVGLPDTVKRCVQQLGLALDFMHGRLVHRDIKPNVLLFDECRVVKLADFGMT 120
 QY 199 RRVGCRVKRVSITPYTAPEVCQAGRADGLAVTGVDFVAFGLVFCVLTGNFPWEAASG 258
 Db 121 RRVGCRVKRVSITPYTAPEVCQAGRADGLAVTGVDFVAFGLVFCVLTGNFPWEAASG 180
 QY 259 ADAFFEEFVWQGRPLPGLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHILT 318
 Db 181 ADAFFEEFVWQGRPLPGLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHILT 240
 QY 319 SELRRRSHRARKPPGDRPPAAGPLRLAEPGLKRTVLTSSGSGSRPAPPVAGSVPLPVP 378
 Db 241 SELRRRSHRARKPPGDRPPAAGPLRLAEPGLKRTVLTSSGSGSRPAPPVAGSVPLPVP 300
 QY 379 VPVPVPVPVPVPVPPGLAPQPGRTDGRADKSGQVVLATAIEICV 424
 Db 301 VPVPVPVPVPVPPGLAPQPGRTDGRADKSGQVVLATAIEICV 346

RESULT 9
 AAU03506
 ID AAU03506 standard; Protein; 572 AA.
 XX
 AC AAU03506;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human protein kinase #6.
 XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200138503-A2.

XX PD 31-MAY-2001.
 XX PF 22-NOV-2000; 2000WO-US32085.
 XX PR 24-NOV-1999; 99US-0167482.
 XX PA (SUGF-) SUGEN INC.
 XX PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 DR WPI: 2001-343950/36.
 DR N-PSDB; AAS06706.
 XX
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Claim 7; Figure 2; 433pp; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 SQ Sequence 572 AA;

Query Match 24.5%; Score 547.5; DB 22; Length 572;
 Best Local Similarity 37.6%; Pred. No. 2.8e-38;
 Matches 130; Conservative 49; Mismatches 134; Indels 33; Gaps 7;
 QY 34 EDMQALTRLTAAADVTYKHYELVRELKGTGTYKVDLVVYKGTGKMAKFPVNSKTKLKN 93
 Db 139 EDMMTLSAQTLVRAEVDLYEVRPLGQGRYGRVLLVTHRQKGTPLALKQLPKPRTSLRG 198
 QY 94 FLREVSITNSLSSTPFIKVFDFWETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKR 153
 Db 199 FLYEFCVGLSLGHAISAVTAYGIGIESAHSYSLTEPVLHGDDLMATIQPKVGUPQAVHR 258
 QY 154 CVQQLGLALDFMHGRLVHRDIKPNVLLFDECRVVKLADFGMTRRVGCRVKRVSITP 213
 Db 259 CAAQLASALEYIHARGLVYRDLKPNVLCDDPACRFKLTDFGHTPRGTLRLAGPPIP 318
 QY 214 YTAPEVC-QAGRADGLAVTGVDFVAFGLVFCVLTGNFPWE-AASGAFAEFEEFVWQR 271
 Db 319 YTAPELCAPPPLPEGLPIQPALDAWALGVLLCLLTGYFPWDRPLAEADPFYEDFLIWA 378
 QY 272 GRIP-GLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHILTSELRRRSHR 330
 Db 379 SGQFRDRPQFWGLAPADALLRGLDDEPHRRSAVIAI----REHLGRFWQREG---- 430
 QY 331 KPFGDRPPAAGPLRLAEPGLKRTVLTSSGSGSRPAPPVAGSVPLP 376
 Db 431 -----EAEAVGVEE----EAGQA-----PAVATEVLP 454

RESULT 10
 ABG30935
 ID ABG30935 standard; Protein; 348 AA.

Db	121	ESCYVFAQEPAGDLFDIIPQVGLFEDTKKRCVQQLGLADFMHSROLVHRDIKPNV	180
QY	181	LLFDRECRVKLADFGMTRRVGCRRKRVSGTIPTAPEVCQAGRADGLAVDTGVDVWAFG	240
Db	181	LLFDRECRVKLADFGMTRRVGCRRKRVSGTIPTAPEVCQAGRADGFAVDITGVDVWAFG	240
QY	241	VLIFCVLTGNTFPWAAAGADAFEEFVFWORGLPGLPSQWRFTPEPALMFORLLALEP	300
Db	241	VLIFCVLTGNTFPWAAAGADAFEEFVFWORGLPGLPSQWRFTPEPALMFORLLALEP	300
QY	301	ERRGPAKSVFRLXHELTSELRRRPSHEARKPPQDRPPAAAGPLRLLEAPGLKRTVLITGSG	360
Db	301	ERRGPAKEVFRLXHELTSELRRRPSHEARKPPQDRLP--GPLRLLEAPGLKRTVLITGSG	358
QY	361	SGSRPAPPAVGSVLPVPVPVVPVVEBPGLAGPGPPGRTDGRADKSKGQVLIATAI	420
Db	359	SGSRSPSPSVGV-----VPVPVPVPVVEBPGLAGPPGRTDGRADKSKGQVLIATAI	413
QY	421	EICV 424	
Db	414	EICV 417	
RESULT 8			
AAU76323			
XX	ID	AAU76323 standard; Protein; 346 AA.	
AC	AAU76323;		
DT	21-MAY-2002	(first entry)	
XX	Human protein kinase 32374.		
XX	Human; protein kinase; 32374; 18431; cancer; infection;		
KW	cellular proliferation; pain; inflammation; metabolic disorder; obesity;		
KW	lipid disorder; diabetes; brain disorder; Alzheimer's disease;		
KW	neural tube defect; cerebrovascular disease; hypoxia; AIDS;		
KW	acquired immunodeficiency syndrome associated myopathy; heart disorder;		
KW	multiple sclerosis; angina pectoris; blood vessel disorder; vasculitis;		
KW	bone disorder; osteoporosis; haematopoietic disorder; autoimmune disease;		
KW	rheumatoid arthritis; osteoarthritis; multiple sclerosis;		
KW	systemic lupus erythematosus; aschma; Grave's disease; liver disorder;		
KW	graft-versus-host disease.		
OS	Homo sapiens.		
XX	WC200210401-A2.		
XX	07-FEB-2002.		
XX	27-JUL-2001; 2001WO-US23653.		
XX	28-JUL-2000; 2000US-221543P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
PA	Meyers R, Kapeller-Libermann R, Silos-Santiago I;		
PI	WPI; 2002-188748/24.		
XX	N-PSDB; ABX10101.		
DR	Novel protein kinase nucleic acid molecules and the encoded proteins		
PT	for diagnosing and treating cellular proliferative, bone, immune,		
PT	cardiovascular, liver, pain or metabolic disorders and identifying		
PT	modulators -		
XX	Claim 4; Fig 1; 141pp; English.		
PS	The invention relates to an isolated protein kinase polypeptide termed as		
XX	32374 or 18431, their allelic variants, fragments and nucleic acids		
CC	encoding them. Also included are a host cell containing the nucleic		
CC	acid, an antibody that binds to the protein kinase and modulators		

CC protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
 CC identifying test compounds, that may act as agonists or antagonists at
 CC the receptor site and which can be regulated to provide therapeutic
 CC effects. The reagent (e.g. antibody or antisense oligonucleotide, which
 CC decreases the expression of human STPKP gene or decreases the levels of
 CC STPKP protein), STPKP agonist, STPKP protein or expression vector is
 CC useful for treating a patient with a chronic obstructive pulmonary
 CC disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
 CC disease, diabetes, obesity or urology disorders. The vectors comprising
 CC STPKP DNA and reagents are useful for preparing a medicament for
 CC modulating the activity of STPKP in the diseases. These diseases include
 CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
 CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
 CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
 CC human STPKP protein.

XX
 SQ Sequence 424 AA;
 Query Match 99.8%; Score 2235; DB 24; Length 424;
 Best Local Similarity 99.8%; Pred. No. 4.6e-183;
 Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHYELVRELG 60
 DB |||||
 QY 1 MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHYELVRELG 60
 DB |||||
 QY 61 KGTYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSSPFFIKVDFVVFET 120
 DB |||||
 QY 61 KGTYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSSPFFIKVDFVVFET 120
 DB |||||
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB |||||
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB |||||
 QY 181 LFDRECRRVKLADFGMTRRVGCRVKSHPGIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
 DB |||||
 QY 181 LFDRECRRVKLADFGMTRRVGCRVKSHPGIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
 DB |||||
 QY 241 VLIFCVLTGNFPWBAASGADAFEEFVRWQGRGLPLPSQWRRTTEPALRMFQLLALEP 300
 DB |||||
 QY 241 VLIFCVLTGNFPWBAASGADAFEEFVRWQGRGLPLPSQWRRTTEPALRMFQLLALEP 300
 DB |||||
 QY 301 ERGPAKEVFRFLKHELTSELRRPSPHRARPPGDRPPAAGPLLEAPGLKRVLTESG 360
 DB |||||
 QY 301 ERGPAKEVFRFLKHELTSELRRPSPHRARPPGDRPPAAGPLLEAPGLKRVLTESG 360
 DB |||||
 QY 361 SGRSPAPAVGSVPLPVPPVPPVPPVPPGLAPGPGPGRTDGRADKSGQVVLATAI 420
 DB |||||
 QY 361 SGRSPAPAVGSVPLPVPPVPPVPPVPPGLAPGPGPGRTDGRADKSGQVVLATAI 420
 DB |||||
 QY 421 EICV 424
 DB 421 EICV 424

RESULT 6

AAE32010

ID AAE32010 standard; Protein; 424 AA.

XX

AC AAE32010;

XX 24-MAR-2003 (first entry)

DE Human serine/threonine protein kinase-like protein (STPKP) #1.

KW Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
 KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
 KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
 KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
 KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
 KW anorectic; neuroprotective; nootropic; cytostatic; uteropathic.

XX Homo sapiens.

OS

XX Key Location/Qualifiers
 PH Domain 53..309
 FT /note= "Eukaryotic protein kinase domain"
 FT Region 59..82
 FT /note= "Protein-kinase-ATP region"
 FT Binding-site 82
 FT /note= "ATP binding site"
 FT Region 170..180
 FT /note= "Protein-kinase-ST region"
 FT Active-site 174
 XX W0200283882-A2.
 XX 24-OCT-2002.
 XX 12-APR-2002; 2002WO-EF04080.
 XX 12-APR-2001; 2001US-283189P.
 PR 04-JUN-2001; 2001US-294989P.
 PR 30-JUL-2001; 2001US-308094P.
 PR 08-FEB-2002; 2002US-354574P.
 XX (FARB) BAYER AG.
 XX Koehler RH;
 XX WPI; 2003-093018/08.
 DR N-PSDB; AAD49416.
 XX
 PT New polynucleotides encoding serine/threonine protein kinase-like
 PT protein (STPKP) and its encoded protein, useful for identifying
 PT modulators of STPKP activity, and in gene therapy for treating e.g.
 PT cancer, obesity or diabetes
 XX
 PS Claim 18; Page 147-149; 161pp; English.
 XX
 CC The invention relates to human serine/threonine protein kinase-like
 CC protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
 CC identifying test compounds, that may act as agonists or antagonists at
 CC the receptor site and which can be regulated to provide therapeutic
 CC effects. The reagent (e.g. antibody or antisense oligonucleotide, which
 CC decreases the expression of human STPKP gene or decreases the levels of
 CC STPKP protein), STPKP agonist, STPKP protein or expression vector is
 CC useful for treating a patient with a chronic obstructive pulmonary
 CC disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
 CC disease, diabetes, obesity or urology disorders. The vectors comprising
 CC STPKP DNA and reagents are useful for preparing a medicament for
 CC modulating the activity of STPKP in the diseases. These diseases include
 CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
 CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
 CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
 CC human STPKP protein.

SQ Sequence 424 AA;

Query Match 98.6%; Score 2207; DB 24; Length 424;

Best Local Similarity 98.6%; Pred. No. 1.2e-180;

Matches 418; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHYELVRELG 60
 DB |||||
 QY 61 KGTYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSSPFFIKVDFVVFET 120
 DB |||||
 QY 61 KGTYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSSPFFIKVDFVVFET 120
 DB |||||
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB |||||
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB |||||
 QY 181 LFDRECRRVKLADFGMTRRVGCRVKSHPGIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240

FT Domain 53..309 /note= "Eukaryotic protein kinase domain"
 FT Domain 54..303
 FT Domain /note= "Protein kinase domain"
 FT Domain 55..304 /note= "Protein kinase domain"
 FT Region 128..141
 FT Region /note= "Tyrosine kinase catalytic site"
 FT Region 164..182
 FT Active-site /note= "Tyrosine kinase catalytic site"
 FT Domain 170..182
 FT Domain 228..248 /note= "Transmembrane domain"
 FT Region 234..256 /note= "Tyrosine kinase catalytic site"
 FT WO200218557-A2.
 XX 07-MAR-2002.
 XX 31-AUG-2001; 2001WO-US27219.
 XX 31-AUG-2000; 2000US-229873P.
 PR 08-SEP-2000; 2000US-231357P.
 PR 14-SEP-2000; 2000US-232654P.
 PR 22-SEP-2000; 2000US-234902P.
 PR 23-SEP-2000; 2000US-236499P.
 PR 06-OCT-2000; 2000US-238389P.
 PR 13-OCT-2000; 2000US-240542P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;
 PI Gururajan R, Ding LS, Patterson C, Yue H, Baughn MR, Tribouley CM;
 PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YI;
 PI Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
 PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
 PI Burford N;
 XX WPI; 2002-329765/36.
 DR N-PSDB; AAD34298.
 XX New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as
 PT leukemia or lymphoma)
 XX Claim 56; Page 141-142; 218pp; English.
 XX The present invention relates to human kinases (PKIN) and polynucleotides
 CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used
 CC in gene therapy and protein therapy. The present sequence is human
 CC PKIN-1 protein.
 XX Sequence 424 AA;
 SQ
 Query Match 99.8%; Score 2235; DB 23; Length 424;
 Best Local Similarity 99.8%; Pred. No. 4.6e-183;
 Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSVCGPEPEPRSTCCGPGAPGAGVPLLTEDMOALTTLTAAADVTWKHYELVRELG 60
 DB 1 MSVCGPEPEPRSTCCGPGAPGAGVPLLTEDMOALTTLTAAADVTWKHYELVRELG 60
 QY 61 KGTYGKVDLVVYKGTGTMALFKVNSKTKLKNFLREVSTNLSLSSPFIKVDVVFET 120

DB 61 KGTYGKVDLVVYKGTGTMALFKVNSKTKLKNFLREVSTNLSLSSPFIKVDVVFET 120
 QY 121 EDCYFAQYAFAGDLFDIIPQVGLPEDIIVKRCVQQLGLALDFMHRGQLVHRDIKPENV 180
 DB 121 EDCYFAQYAFAGDLFDIIPQVGLPEDIIVKRCVQQLGLALDFMHRGQLVHRDIKPENV 180
 QY 181 LLDRECRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCAQAGRADGLAVDTGVDVWAFG 240
 DB 181 LLDRECRVKLADFGMTTRVRCRVKVSQGTIPYTAPEVCAQAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFWEAASGADAFEEBFVVRWQGRGLPGLPSQWRRTFEPALRMFORLLALEP 300
 DB 241 VLIFCVLTGNFWEAASGADAFEEBFVVRWQGRGLPGLPSQWRRTFEPALRMFORLLALEP 300
 QY 301 ERGPAKEVFRFLKHELTSELRRRPSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 DB 301 ERGPAKEVFRFLKHELTSELRRRPSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 QY 361 SGSRPAPPAVGSVPLPVVPVVPVPEPLAQGGPGRDGRADSKSGQVLTALAI 420
 DB 361 SGSRPAPPAVGSVPLPVVPVVPVPEPLAQGGPGRDGRADSKSGQVLTALAI 420
 QY 421 EICV 424
 DB 421 EICV 424
 RESULT 5
 AAE32011
 ID AAE32011 standard; Protein; 424 AA.
 XX AAE32011;
 AC AAE32011;
 XX 24-MAR-2003 (first entry)
 DT Human serine/threonine protein kinase-like protein (STPKP) #2.
 DE Human
 XX Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
 KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
 KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
 KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
 KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hyperplasia;
 KW anorectic; neuroprotective; nootropic; cytosstatic; uropathic.
 XX Homo sapiens.
 OS WO200283982-A2.
 PN 24-OCT-2002.
 PD 12-APR-2002; 2002WO-EF04080.
 PF 12-APR-2001; 2001US-283189P.
 PR 04-JUN-2001; 2001US-294989P.
 PR 30-JUL-2001; 2001US-308094P.
 PR 08-FEB-2002; 2002US-354574P.
 XX (FARB) BAYER AG.
 PA Koehler RH;
 PI WPI; 2003-093018/08.
 DR N-PSDB; AAD49419.
 XX New polynucleotides encoding serine/threonine protein kinase-like
 PT protein (STPKP) and its encoded protein, useful for identifying
 PT modulators of STPKP activity, and in gene therapy for treating e.g.
 PT cancer, obesity or diabetes
 XX Claim 18; Page 157-159; 161pp; English.
 XX The invention relates to human serine/threonine protein kinase-like

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:50:21 ; Search time 47 Seconds
(without alignments)
1431.916 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPPPPRLTCCGPG.....GRADKSKGVVLTATAIEICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	23 AAU79585	Novel human protei
2	2239	100.0	424	23 AAU10023	Human protein kina
3	2235	99.8	424	22 AAU03544	Human protein kina
4	2235	99.8	424	23 AAU21706	Human PKIN-1 prote
5	2235	99.8	424	24 AAU32011	Human serine/threo
6	2207	98.6	424	24 AAU32010	Human serine/threo
7	2084.5	93.1	417	21 AAU83904	Rat PKs protein.
8	1822	81.4	346	23 AAU76323	Human protein kina
9	547.5	24.5	572	22 AAU03506	Human protein kina

10	546.5	24.4	348	23	ABG3093E	Human novel protei
11	546.5	24.4	348	23	AAE21721	Human PKIN-16 prot
12	508.5	22.7	373	22	AAU03507	Human protein kina
13	508.5	22.7	414	24	AAE32027	Human kinase and p
14	452.5	20.2	577	22	ABE61118	Drosophila melanog
15	438	19.6	456	22	ABE65543	Drosophila melanog
16	383.5	17.1	250	23	ABE78796	Serine/threonine k
17	367	16.4	761	24	ABP96068	Human protein kina
18	361	16.1	702	22	ABP70124	Drosophila melanog
19	355	15.9	366	24	ABP96083	Human protein kina
20	352	15.7	520	22	AAE65629	Novel protein kina
21	351.5	15.7	686	22	AAE65621	Novel protein kina
22	350.5	15.7	825	24	ABP97380	Human DCAMKL1-like
23	350	15.6	254	23	ABE78797	Protein kinase dom
24	349.5	15.6	252	23	ABE78821	Protein kinase dom
25	347.5	15.5	724	23	ABE04431	Murine neuronal se
26	346	15.5	817	23	AAO15419	Novel human kinase
27	346	15.5	835	23	ABE82471	Human serine/threo
28	346	15.5	835	23	AAE24133	Human kinase (PKIN
29	345.5	15.4	514	22	ABE60741	Drosophila melanog
30	345.5	15.4	514	22	ABE60742	Drosophila melanog
31	345.5	15.4	660	23	ABE04432	Murine neuronal se
32	344	15.4	512	21	AAE36157	Arabidopsis thalia
33	343.5	15.3	509	21	ABE03416	Corn putative carb
34	343.5	15.3	509	24	ABE40708	Zea mays oil trait
35	343.5	15.3	509	24	ABE40809	Zea mays oil trait
36	343	15.3	512	23	AAO17663	Oryza sativa oil t
37	342.5	15.3	531	24	ABE40718	A thaliana AKIN11.
38	342	15.3	512	23	AAO17664	Triticum aestivum
39	342	15.3	523	21	AAE03425	A thaliana AKIN11
40	342	15.3	523	24	ABE40719	Wheat putative car
41	341.5	15.3	688	23	ABE04434	Triticum aestivum
42	341.5	15.3	729	19	AAE37158	Human neuronal ser
43	341.5	15.3	729	22	ABE65628	Human Twenty-five
44	341.5	15.3	729	24	AAE33552	Novel protein kina
45	341.5	15.3	752	22	AAE11782	Human microtubule
						Human kinase (PKIN

ALIGNMENTS

RESULT 1
AAU79585
ID AAU79585 standard; Protein; 424 AA.
AC AAU79585;
XX
XX 24-SEP-2002 (first entry)
XX
XX Novel human protein (NHP).
DE
XX Human; Novel human protein; NHP; kinase; regulatory pathway;
KW therapeutic; diagnostic; pharmacogenomic; antigenic.
XX
XX Homo sapiens.
XX
XX WO200246428-A2.
XX
XX 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US48533.
XX
XX 07-DEC-2000; 2000US-251941P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Maricar M, Yu X, Fridele CJ;
XX
XX WPI; 2002-527921/56.
XX
XX N-PSDB; ABK86975.
XX
XX Novel nucleic acid molecule encoding a human kinase, useful in
XX therapeutic, diagnostic and pharmacogenomic applications, as DNA
XX

RT seed development.";
 RL Mol. Gen. Genet. 260:388-394 (1998).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; D82038; BAA36297.1; -.
 DR EMBL; D82036; BAA36295.1; -.
 DR HSSP; Q63450; 1A06.
 DR Gramene; Q92NT4; -.
 DR InterPro; IPR001772; Kinase Cterm.
 DR InterPro; IPR000715; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF02149; KAI; 1.
 DR Pfam; PF00669; pkinase; 1.
 DR Pfam; PF00627; UBA; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_St; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 508 AA; 58250 MW; E935262080B39A59 CRC64;

Query Match 15.3%; Score 343.5; DB 10; Length 508;
 Best Local Similarity 32.2%; Pred. No. 1.7e-19;
 Matches 95; Conservative 60; Mismatches 109; Indels 31; Gaps 11;
 QY 47 SDVTKHVELVRELKGYGVKVDLVVYKGTGKVALKFNKSKTKLKNF-----LREVS I 100
 Db 11 SEALKNYNLGRTLGIGSGFKVIAEHKLTGHRVAIKILN--RRQMRNMEMEEKAKREIKI 68
 QY 101 TNSLSSSPFIKVDVVFETEDCVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQOLGL 160
 Db 69 LR-LFIHPHILRIUYEVITYPTDIYV-VMEYCKFGELFDYVEKGRLOEDARRIFQOIIIS 126
 QY 161 ALDFMHGRQLVHRDIKPEENVLLFDRECRVKLADFGMTTRV--GCRVKRVSGTIPYTAPE 218
 Db 127 GVEYCHRNVMVHRDLKPENLDDSK--YNYKLADFGLSNVMDHGHFLKTCGSPNYAAPE 184
 QY 219 VCQAGRADGLAVDTGVDFWAFGLVIFCVLTGNFPWEAASGADAFEEFVWQGRPLGLP 278
 Db 185 VI-SGK---LYAGEVDVWVSCGVILYALCGTLFPD-----DENIPNLFKKIKGGIYTL 235
 QY 279 SQWRRTFEPALRMFQRLALPEPRGPAKEVFRFLKHELTSELRRPSHARKPP 333
 Db 236 S----HLSDALDLIPRLVVDPMKRTITREIRHQWFQI-----RLPRYLAVPPP 282

RESULT 15
 Q92RJ3 PRELIMINARY; PRT; 509 AA.
 AC Q92RJ3
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE OSK4.
 GN OSK4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99086251; PubMed=980704;
 RA Takano M., Kajiya-Kanegae H., Funatsuki H., Kikuchi S.;
 RT "Rice has two distinct classes of protein kinase genes related to SNF1
 of Saccharomyces cerevisiae, which are differently regulated in early
 seed development.";
 RT Mol. Gen. Genet. 260:388-394 (1998).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; D82035; BAA36299.1; -.
 DR

DR HSSP; Q63450; 1A06.
 DR Gramene; Q92RJ3; -.
 DR InterPro; IPR001772; Kinase Cterm.
 DR InterPro; IPR000715; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF02149; KAI; 1.
 DR Pfam; PF00669; pkinase; 1.
 DR Pfam; PF00627; UBA; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_St; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 509 AA; 58471 MW; 6E82566D715A1B48 CRC64;

Query Match 15.3%; Score 343.5; DB 10; Length 509;
 Best Local Similarity 32.2%; Pred. No. 1.7e-19;
 Matches 95; Conservative 60; Mismatches 109; Indels 31; Gaps 11;
 QY 47 SDVTKHVELVRELKGYGVKVDLVVYKGTGKVALKFNKSKTKLKNF-----LREVS I 100
 Db 11 SEALKNYNLGRTLGIGSGFKVIAEHKLTGHRVAIKILN--RRQMRNMEMEEKAKREIKI 68
 QY 101 TNSLSSSPFIKVDVVFETEDCVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQOLGL 160
 Db 69 LR-LFIHPHILRIUYEVITYPTDIYV-VMEYCKFGELFDYVEKGRLOEDARRIFQOIIIS 126
 QY 161 ALDFMHGRQLVHRDIKPEENVLLFDRECRVKLADFGMTTRV--GCRVKRVSGTIPYTAPE 218
 Db 127 GVEYCHRNVMVHRDLKPENLDDSK--YNYKLADFGLSNVMDHGHFLKTCGSPNYAAPE 184
 QY 219 VCQAGRADGLAVDTGVDFWAFGLVIFCVLTGNFPWEAASGADAFEEFVWQGRPLGLP 278
 Db 185 VI-SGK---LYAGEVDVWVSCGVILYALCGTLFPD-----DENIPNLFKKIKGGIYTL 235
 QY 279 SQWRRTFEPALRMFQRLALPEPRGPAKEVFRFLKHELTSELRRPSHARKPP 333
 Db 236 S----HLSDALDLIPRLVVDPMKRTITREIRHQWFQI-----RLPRYLAVPPP 282

Search completed: December 5, 2003, 09:57:44
 Job time : 41 secs

```

Db 525 AAGVILYILLCGPPPPSPERQDELFINI---QVGQFFELSPYMDNISDAAKDLVRNLL 581
QY 297 ALEPERGPAKVFRLKHELTSELRRRSHRARKP 332
Db 582 EVDPKRYTAEQLQHPWIEVGHNTGNSQKEESP 617

RESULT 13
P92958
AC P92958 PRELIMINARY; PRT; 512 AA.
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase (AKIN11 protein kinase) (AT3g29160/MXE2_16).
GN AKIN11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Bhalerao R.P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Nakamura Y.;
RL MEDLINE=20277480; PubMed=10819329;
RX
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
[4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Huan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X99279; CAB67671.1; -
DR EMBL; AB018121; BAB01933.1; -
DR EMBL; AY070468; AAL49934.1; -
DR EMBL; AY149927; AAN31081.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot_kinase.

```

```

DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 512 AA; 58689 MW; 4FBFE5D3EC60F87 CRC64;

Query Match 15.4%; Score 344; DB 10; Length 512;
Best Local Similarity 28.5%; Pred. No. 1.5e-19;
Matches 120; Conservative 65; Mismatches 132; Indels 104; Gaps 20;

QY 49 VTKHYELVRELKGYGVKVDLVYKGTGTMALKFVNKSKTK-----LKNFLREVSIITNSL 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 16 ILPNYKLGKTLGIGSGFKVIAEHVVTGHKVAIKILNRRKIKNMWEEKVRSEIKILR-L 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 SSSPFIKVFVDVVFETEDCYFAQVAPAGDLFDIIPQVGLPDTVKRCVOQLGLALDF 164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 FMHPHIIQYEVITTSDIYV-VMEYVKSGLFDVIVEKRLQDEARNFFQOIIISGVET 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 MEGRQLVHRDIKPNVLLFDRECRVKLADFQWTR--RVGCRVKRVSGTIPYTAPEVCOA 222
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 CHRNWVHRDLKPNLLL--DSEC-NIKLADFGLSNVMDRGHLTKTSCGSPNYAAPEVI-S 190
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 GRADGLAVDTGVDVNAFVGLIFCVLTGPNPWAASGADAFFEEFVWQGRPLPGLPSQW- 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 GK---LYAGPEVDVMSGCVILYALLCGTLFPD-----DENIPLFKIKGGIYTLPSHLS 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 282 -----RRTEPALRM---FQ-----RLLALE-PERRGPAKEV----- 309
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 SEARDLIPRLMLIVDPVKRITIPRIQHWFWQTHLPYLAISPDPDTVQAKKNEEIVQEV 302
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 --FRFLKHELTSELRRRP-----SHRAKPKPGDRPPAAGPLRLAEPGLKRTV 355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 VNMGFDNQVLESRLNRNTQNDATVTYVLLDNRFRVPSG-----YLES--EFQET- 350
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 LTESGSGSRPAPVAGSVLPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 -TDSGN-----PMRTPEAGASPVGHWTIPAHVDHYGLGARSQ 386
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 V 414
Db 387 V 387

RESULT 14
Q9ZNT4 PRELIMINARY; PRT; 508 AA.
ID Q9ZNT4
AC Q9ZNT4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSK3.
GN OSK3 OR OSK5. (Rice).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99086251; PubMed=9870704;
RA Takano M., Kajiya-Kanegae H., Funatsuki H., Kikuchi S.;
RT "Rice has two distinct classes of protein kinase genes related to SNF1
RT of Saccharomyces cerevisiae, which are differently regulated in early

```


RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.M., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
CC Science 287:2185-2195(2000).
CL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB003798; AAF57652.1; -;
DR HSP; Q63450; 1A06.
DR FlyBase; FBgn034376; CG15072.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 702 AA; 77356 MW; 441253E5A5484647D CRC64;

Query Match 16.1%; Score 361; DB 5; Length 702;
Best Local Similarity 29.1%; Pred. No. 9,7e-21;
Matches 109; Conservative 72; Mismatches 148; Indels 46; Gaps 15;

QY 18 GPGTAPGGAGVPLLTEDMQALILRLAASDVTK-----HYELVRELKGYTKGVLDVYV 72
DB 8 GPAAP-PTSSTP-----QNYKVPSTKISVDKLLRGVGYELEKTIKGNFAVVKLATN 60
QY 73 KGTGTMALKFVNK---SKTKLKNFLREVSTINSLSSSPFIKVFDDVVFETDCYVFAQE 129
DB 61 IVYKTKVAIKIDKTCNBEYLNKTFREIALKSL-RHPHITRLIYVYM-ESQNIYLVTE 118
QY 130 YAPAGLDFDIIPQGVLPEDTVKRCVQOGLAUMHGRQLVHRDIKPNVLLFDRECR 189
DB 119 YAPNGEIFDLVANGRMKEPEARVFTQLVSAVHYCHRRGVVHRDLKAENVLL-DKD-MN 176
QY 190 VKLADFGWTRRV--GCRVKRVSGLTIPVTAPEVCOAGRADGLAVDTGVDVWAFGLIFCVL 247
DB 177 IKLADFGSNHYERGATIKTCWGSPPYAAPEVFGLEYDG-----PKSDIWSLGVVLYALV 232
QY 248 TGNFPFWEAASGADAFFEEFVRWQRGLPGLPSPQWRREFTEPALRMFORLLALEPERRGPAK 307
DB 233 CGALPFDPGKITLE--LKSRLVIGKRFIPFNSQ-----ECEQLIRNMLVVEPDRRTIK 284
QY 308 EVFELKHELTSELRPRSHARKPPGDRPPAAGPLRLAPGPKLVITLSSGSGSPAP 367
DB 285 QI-----IKHRMLSEQSEMQEERF--GDMSPGS-----GTVSKSASTSSLSGASDSP 331
QY 368 PAVGSVPLPVPVVPV 382
DB 332 PQLDSVNVVTHMLQLP 346

RESULT 9
Q960N7 PRELIMINARY; PRT; 512 AA.
ID Q960N7
AC Q960N7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LD42896p.

GN GRP OR CG4711 OR CG17161.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RP STRAIN=Berkely;
RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacled J., Paragas V., Park S., Prounanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AY051961; AAK93385.1; -;
DR FlyBase; FBgn0011598; grp_kinase.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 512 AA; 57833 MW; 504CAE3C0B642EAO CRC64;

Query Match 15.9%; Score 356.5; DB 5; Length 512;
Best Local Similarity 32.7%; Pred. No. 1.5e-20;
Matches 87; Conservative 49; Mismatches 111; Indels 19; Gaps 7;

QY 45 AASDVTKHYELVRELKGYTKGVLDVWYKGTGTPMALKFVN-KSKTKLKNFLREVSTINS 103
DB 14 ATREFVEGVTLAQTLGEGAYGEVKLLINRQTGEAVAMKMWDLKHPDAANSVRKEVCIOK 73
QY 104 LSSSPFIKVFDDVVFETDCYVFAQYAPAGLDFDIIPQGVLPEDTVKRCVQOGLALD 163
DB 74 MLQDRHILFFGKRSQGSVEYIFL-EYAAAGGELFDRIEPDVGMFQPEAQRIFTQLLSGLN 132
QY 164 FMHGROLVRHDIKPNVLLFDRECRVVKLADFGWTRRVGCRVK-----RVSGTIPVTAPE 218
DB 133 YLHQGIARHDLKPNLLDDEHD--NVKISDFGMATMFRCKGKERLLDKCGTLPIVAPE 190
QY 219 VQO-AGRADGLAVDTGVDVWAFGLIFCVLTGTFPWEAASGADAFFEEFVRWQRGLPGL 277
DB 191 VLQKAYHAQ-----PADLWSCGVILVTLAGELPMDQPS---TNCTEFTNWRDNDHWQL 241
QY 278 PSQWRREFTEPALRMFORLLALEPERR 303
DB 242 QTFWSKLDTLATSLRLKLLATSPGTR 267

RESULT 10
Q9C098 PRELIMINARY; PRT; 608 AA.
ID Q9C098
AC Q9C098
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1765 (Fragment).
GN KIAA1765.
OS *Homo sapiens* (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP MEDLINE=21082932; PubMed=11214970;
RX Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code

```
DR EMBL: AY071597; AAL49219.1; -
DR FlyBase; FBgn0031855; CGI1221.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 456 AA; 52418 MW; 272D73C6677C375B CRC64;

Query Match 19.6%; Score 438; DB 5; Length 456;
Best Local Similarity 34.1%; Pred. No. 3e-27;
Matches 113; Conservative 54; Mismatches 132; Indels 32; Gaps 8;

QY 5 CPPEPPRSRLTCCGPGTAGPGA-----GVPLLTDMQALTLRTLAASD 48
Db 44 CSTPLTFS-TSTGRLEVPGAASVRSSTYKKPKNDGGQIHLIPDVE-LPLMTFA--- 98
QY 49 VTKHYELVRELKGTGKGVKLVVYKGTGKMAKLVFNKSKTKLKNFLREVSITNSLSSSP 108
Db 99 --DQYNIEKTLAEGCFACKILLCRHRPNTLVLKAVHAELTTIKFQKEFHYNYELSHH 156
QY 109 FLIKVDFVETEDCYFAQYAPAGLDLFIIPQVGLPDTVKRCVQQLGLALDFMHGR 168
Db 157 HILSAYAVAFQTMQYVYFAMHAFYGLASNGN-GLHENACKLISEQLSALLGFVHSH 215
QY 169 QLVHRDLKPNVLLFDRECRVKLADFGMTRRVRCRVKVSQT---IPYAPVQCAGR 224
Db 216 NLVHRDLKINILVFTDFRVKLCDFGATTKGILLVHKVHTWTSVCPPEQLELIKNER 275
QY 225 ADGLAVDTGVDVNAFGLVIFCVLTGNFPWEAASGA-DAPFEFVRWQGRGLPGLPSQMR 283
Db 276 FQCLPVS---DSWQFGILLYNLITGNPPQSDWVDKQSYANFMKYEQRTTKVDPNFR 332
QY 284 FTPEALRMFORLLALEPBERGPAKEVFRFLK 314
Db 333 FSPRLMRCFRKYLSDPEDRCRKITEVAKYMK 363

RESULT 7
Q90ZY5 PRELIMINARY; PRT; 503 AA.
AC Q90ZY5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein_kinase Chk2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
R1 Chou C.-M., Leu J.-H., Huang C.-J.;
RT "The zebrafish homolog of the human checkpoint kinase, Chk2."
RC Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF265346; AAK52419.1; -
DR InterPro; IPR000253; FHA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00498; FHA; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 503 AA; 56699 MW; EC01BF3D8E4CEC71 CRC64;

Query Match 16.5%; Score 370; DB 13; Length 503;
Best Local Similarity 30.6%; Pred. No. 1.2e-21;
Matches 107; Conservative 57; Mismatches 148; Indels 38; Gaps 13;

QY 32 LTEDMQALTLTLAASDVTKHYELVRELKGTGKGVKLVVYKGTGKMAKLVFNKSK--- 88
Db 178 MADDQANLPL-----EFSKKYHIAKIGTGVGCVKLAIEKETFKKVALTKINKHDFPS 231
QY 89 --TKLKNFLREVSIITNSLSSPFIILKVFVVFTEDCYFAQYAPAGLDLFIIPQVGL 146
Db 232 IGATRNAERIEILKKI-DHPCLIKTED-FYOTDSYIVLYIEGGELEFGRIKAKKKL 289
QY 147 PEDITVKRCVQQLGLALDFMHGROLVHRDIKPNVLL--FDRECRVKLADFGMTR--RVG 202
Db 290 EEDIKLYFYQMLKAVLYLHNGIILHRLKPNVLLASHDDIC-LIKITDFNQSKILEES 348
QY 203 CRVRSVGTITPYTAPVCCQAGRADGLAVDTGVVNAFGLVIFCVLTGNFPWEAASGADF 262
Db 349 SLMKTLICGTPTYVAPEVFT--HASTVGYTKADVWSLGLVLLFTCLGGYPPEPTECTTMSV 406
QY 263 FEFSEVWQGRGLPGLPSQMRRETEPALRMFORLLALEPBERGPAKEVFRFLKHELTSER 322
Db 407 REGI---NGEYFIISQWKNVSEAKDLVKLLVDPQKRLSVEEA---LEHPWLKDDR 460
QY 323 RPSHRARKEPGRPPAAGPLRLAEPGLKRTVLTESGSGS----RPAP 367
Db 461 MRQANQLMNPQ---AANQPMRPEATIRK---AQEGEGEPSSKRKPGP 503

RESULT 8
Q9V8L2 PRELIMINARY; PRT; 702 AA.
AC Q9V8L2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG15072 protein.
GN CG15072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
R1 MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers V.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Aghayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Shandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegun C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```



```
Best Local Similarity 64.8%; Pred. No. 7.7e-74;
Matches 184; Conservative 43; Mismatches 57; Indels 0; Gaps 0;

Qy 32 LTEMQALTLRTAASDVTHYELVRELKGTGKVDLVVYKGTGTMALKFVNKSTKL 91
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 11 ILBELQYTAQNLKLEVNKYEVIRELKGTYKVDLVHKGSMALKFLKSKSTKL 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 92 KNFLREVSITNSLSSFFIIKVFDFPTEPCYVFAQYAPAGDLFIIPQVGLPEDTV 151
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 71 KSFLREYSISLYSPCFEINMGIAFETDHYVFAQYAPSGDLFIIPQVGLPEPVA 130
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 152 KRCVQOGLALDFMHGQLVHRDIPKENVLLFDRECRVKLADFGMTTRVGCVRKVS 211
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 131 KRCVHQVAIALEYLHSLKVRDIPKENVLLFDRECRVKLADFGMTTRVGCVRKVS 190
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 212 IPYTAPEVCQAGRADGLAVDGVWAFGLVFCVLTGNFWEAASGADAFEEFVRWQ 271
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 191 IPYTAPELCTSKHDGFCVDYSTDVWAFGLVFCVLTGNFWEAASGADAFEEFVR 250
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 272 GRPLGSLQWRFTPEALRMPQRLALEPERRGPAKEVFRFLKH 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 251 RRTGAVSQWRFTDESILRFRKLLALEQERRCSVKVEFAHLGH 294
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q91758 PRELIMINARY; PRT; 358 AA.
AC Q91758;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97042347; PubMed=8887547;
RA Snape A.M., Smith J.C.;
RT "Regulation of embryonic cell division by a Xenopus gastrula-specific
protein kinase.";
RL EMO J. 15:4556-4565 (1996).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X99406; CAA67783.1; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 358 AA; 40857 MW; 300086D06F59837B CRC64;

Query Match
Best Local Similarity 28.7%; Score 642.5; DB 13; Length 358;
Matches 125; Conservative 46; Mismatches 95; Indels 3; Gaps 2;

Qy 47 SDVTKHYELVRELKGTGKVDLVVYKGTGTMALKFVNKSTKLKFLREVSITNSLSS 106
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 5 SOIGDHYHVIKLGQTFSHVLLARELISGHSVALKVRKDXRTQAAPFIHNLISLSD 64
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 107 SPFIKVFDDVFETEDCVFAQYAPAGDLFIIPQVGLPEDTVKRCVQVGLALDFMH 166
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 65 YPGIKTYPTVETVDVFIPTQELAPATLSIIKTEVIGFEEVYVRCAVQITALDFMH 124
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 167 GRQLVHRDIPKENVLLFDRECRVKLADFGMTTRVGCVRKVSQTIPTAPEVCAQGRAD 226
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 125 GRRLVHRDIPKENVLLMDKCYHKLCDLFGTQSVGLVSPNSHIPIPMPELNLKPNQ 184
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Qy 227 GLAVDTGVDVWAFGLVFCVLTGNFWEAASGADAFEEFVRWQGR--LPGLPSQWRRF 284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 185 LVLVDQSDIWSGILLFVALTGYFFWEAVGHQKQYQMFVHQNNNYPA-PILANRF 243
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 285 TEPALRMPQRLALEPERRGPAKEVFRFL 313
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 244 TQEMAMFFKLKLSQKFSRSPDPTVLEFL 272
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
Q8SZT8 PRELIMINARY; PRT; 551 AA.
AC Q8SZT8; Q9V7R1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH19382p (CG4945 protein).
GN CG4945.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarino H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RN [2]
RP Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.Y.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwack J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Bolland J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
```


[illegible]

```

RESULT 2
Q8QZX0 PRELIMINARY; PRT; 417 AA.
AC Q8QZX0;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SBK.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
EMBL; BC024114; AAH24114.1; -.
EMBL; BC025837; AAH25837.1; -.
EMBL; BC031759; AAH31759.1; -.
DR MGD; MGI:2135937; Sbk.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;

```

KW	Serine/threonine-protein kinase; Transferase.
SQ	SEQUENCE 417 AA; 45696 MW; C6A79A28EE6D2EC7 CRC64;
	Query Match 92.6%; Score 2072.5; DB 11; Length 417;
	Best Local Similarity 94.1%; Pred. No. 3.7e-160;
	Matches 399; Conservative 3; Mismatches 15; Indels 7; Gaps 2;
QY	1 MSVGCPEPPERSLTCGPGTAPGAGVPGLLEDMQALTRTLAAADVTVKHYELVRELG 60
Db	1 MSVGCPEPPHSLPCCGPGAAPVGAVGPLLTEDMQALTRTLAAASDVTKHYELVRELG 60
QY	61 KGTYGKVDLVVYKGTGTHAKLKFNKSKTKLNFLREVSTNSLSPPFIKKVFDDVVFFET 120
Db	61 KGTYGKVDLVAYKGTGTHAKLKFNKSKTKLNFLREVSTNSLSPPFIKKVFDDVVFFET 120
QY	121 EDCVFAEQYAPAGDLFDIIPPOGVLPEDTVKRCVOQLGLALDFMHGRQLVHRDIKPENV 180
Db	121 EECVFAEQYAPAGDLFDIIPPOGVLPEDIVKRCVOQLGLALDFMHSRQLVHRDIKPENV 180
QY	181 LLEFDECRVRKLADFNGTRRVGCRVRKVSCTIPTYAPEVCQAGRADGLAVDTGVDVMAFG 240
Db	181 LLEFDECRVRKLADFNGTRRVGCRVKVSCTIPTYAPEVCQAGRADGFVDTGVDVMAFG 240
QY	241 VLIFCVLITGNFPWEAASGADAFFEFEVWGRLPGIPSGWRBFTTEPALRMFORLLALEP 300
Db	241 VLIFCVLITGNFPWEAASGADAFFEFEVWGRLPLGPSQWRRTTEPALRMFORLLALEP 300
QY	301 ERGGPAKEVFRFLKHELTSRLRRPSHFARKPPQDRPPAAPGLRLAEPGLKRTVLTESG 360
Db	301 ERGGPAKEVFRFLKHELTSRLRRPSHFARKPPQDRPL--GSIRLEAPGLKRTVLTESG 358
QY	361 SGRSPAPPVAGSVPLPVPPVPVPPVEPGLAPOGPGRDTGRADKSGQVVLATAI 420
Db	359 SGRSPSPSPSGPV-----VPVPVPVPEAGLAPPAPGPRDTGRDKSGQVVLATAI 413
QY	421 EICV 424
Db	414 EICV 417
RESULT 3	
Q90ZY4	PRELIMINARY; PRT; 385 AA.
ID	Q90ZY4
AC	Q90ZY4; 2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Protein Kinase Bsk146.
OS	Brachydanio rerio (zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RA	SEQUENCE FROM N.A.
RP	Chou C.-M., Leu J.-H., Huang C.-J.;
RT	"Molecular cloning and characterization of a novel protein kinase
RT	expressed predominantly in the brain from zebrafish.";
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR	EMBL; AF265347; AAK52420.1; -
DR	InterPro; IPR000719; Prot.kinase.
DR	InterPro; IPR002290; Ser.thr.pkinase.
DR	InterPro; IPR001245; Tyr.pkinase.
DR	Pfam; PF00069; pkinase; 1.
DR	PRINTS; PR00109; TYRKINASE.
DR	PRODOM; PD000001; Prot.kinase; 1.
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ	SEQUENCE 385 AA; 43913 MW; 54BE4E6D3819EBD2 CRC64;
	Query Match 45.1%; Score 1010; DB 13; Length 385;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:54:06 ; Search time 40 Seconds
(without alignments)
2735.358 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGPPEPRSLTCGPG.....GRADSKQGVLAIAICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2084.5	93.1	417	11 Q9Z335	Q9Z335 rattus norv
2	2072.5	92.6	417	11 Q8QZX0	Q8QZX0 mus musculus
3	1010.1	45.1	385	13 Q90ZY4	Q90ZY4 brachydanio
4	642.5	28.7	358	13 Q91758	Q91758 xenopus lae
5	452.5	20.2	551	5 Q8SZT8	Q8SZT8 drosophila
6	438.1	19.6	456	5 Q9VM90	Q9VM90 drosophila
7	370.1	16.5	503	13 Q90ZY5	Q90ZY5 brachydanio
8	361.1	16.1	702	5 Q9V8L2	Q9V8L2 drosophila
9	356.5	15.9	512	5 Q960N7	Q960N7 drosophila
10	346.1	15.5	608	4 Q9C098	Q9C098 homo sapien
11	345.5	15.4	752	11 Q8C1P4	Q8C1P4 mus musculus
12	344.5	15.4	619	11 Q8BQ05	Q8BQ05 mus musculus
13	344.1	15.4	512	10 P92958	P92958 arabidopsis
14	343.5	15.3	508	10 Q9ZNT4	Q9ZNT4 oryza sativ
15	343.5	15.3	509	10 Q9ZRX3	Q9ZRX3 oryza sativ
16	343.1	15.3	476	13 Q8AYC9	Q8AYC9 gallus gall

17	343	15.3	512	10 P92968	P92968 arabidopsis
18	342.5	15.3	485	5 Q9N3L4	Q9N3L4 caenorhabdi
19	340.5	15.2	640	11 Q8C0X8	Q8C0X8 mus musculus
20	340.5	15.2	729	11 Q9JKE4	Q9JKE4 mus musculus
21	340.5	15.2	744	11 Q9JKE5	Q9JKE5 mus musculus
22	340.1	15.2	513	5 Q81661	Q81661 drosophila
23	339.5	15.2	716	4 Q8N399	Q8N399 homo sapien
24	336.5	15.0	535	10 Q8RWD2	Q8RWD2 arabidopsis
25	336.5	15.0	752	4 Q8NG37	Q8NG37 homo sapien
26	336.5	15.0	797	11 Q8VHF0	Q8VHF0 rattus norv
27	336.1	15.0	504	10 P93113	P93113 cucumis sat
28	336.1	15.0	795	4 Q9P0L2	Q9P0L2 homo sapien
29	335.1	15.0	474	13 Q9Y1I8	Q9Y1I8 xenopus lae
30	334.5	14.9	517	13 Q98TW0	Q98TW0 xenopus lae
31	334.5	14.9	517	13 Q918V3	Q918V3 xenopus lae
32	334.1	14.9	778	4 Q8TDC2	Q8TDC2 homo sapien
33	334.1	14.9	793	11 Q08678	Q08678 rattus norv
34	334.1	14.9	795	11 Q8VHJ5	Q8VHJ5 mus musculus
35	333.1	14.9	476	11 Q9DON2	Q9DON2 mus musculus
36	332.5	14.8	504	10 Q41485	Q41485 solanum tub
37	332.1	14.8	510	10 Q40740	Q40740 oryza sativ
38	332.1	14.8	511	10 Q40544	Q40544 nicotiana t
39	332.1	14.8	513	10 Q40029	Q40029 hordeum vul
40	331.5	14.8	328	5 Q9NF25	Q9NF25 caenorhabdi
41	330.5	14.8	476	11 Q91ZN7	Q91ZN7 rattus norv
42	330.5	14.8	505	10 Q9ZRX1	Q9ZRX1 oryza sativ
43	330.1	14.7	514	10 Q9M726	Q9M726 lycopersico
44	329.5	14.7	291	10 Q9ZRA0	Q9ZRA0 arabidopsis
45	329.5	14.7	494	10 Q9FLZ3	Q9FLZ3 arabidopsis

ALIGNMENTS

RESULT 1

Q9Z335 ID Q9Z335 PRELIMINARY; PRT; 417 AA.
AC Q9Z335;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Serine/threonine protein kinase (EC 2.7.1.-).
GN SBK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WistarST; TISSUE=Hippocampus;
RA Nara K.;
RT "Cloning and characterization of a novel serine/threonine protein
RT kinase gene expressed dominantly in developing brain.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB010154; BAA36362.1; -;
DR HSSP; Q00534; 1B18.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 417 AA; 45676 MW; C4D36608EE6EE208 CRC64;

Query Match 93.1%; Score 2084.5; DB 11; Length 417;
Best Local Similarity 94.6%; Pred. No. 3.9e-161;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

Matches	95; Conservative	60; Mismatches	109; Indels	32; Gaps	12; Gaps
QY	47	SDVTKHYELVRELGGKTYGKVDLVVYKGTGTQWALXPNKSKTKLKNF-----LRVSI	100		
Db	11	SEALKNYNLGRTLGIGSGFKVKAETHLGTGRVAIKILN--RRQMRNMEEEKAKREIKI	68		
QY	101	TNSLSSSPFIKVDVVFFETEDCVPAQAEYAPAGDLFDIIPVQGLPEDTVK--RCVOOLG	159		
Db	69	LR-LFTHPHILRIEYVIYPTDIYV-VMEXCKEGELFDYIVKGRLOEDRVALRIFSQII	126		
QY	160	LALDFMHGQQLVHRDIKPNVLLDFEGRVVKLADFGMTRRV--GCRVKRVSGTPIYTPAP	217		
Db	127	SAVEYCHRNVAHRDLKPNLLDSK--YNVKLADFGLSNMWHDGHFLKTKSGSPNTAAAP	184		
QY	218	EVYQAGRADGLAVDTGVDVWAFGVLIFCVLTGNFPMEASGADAFEEFVVRWGRGLPGL	277		
Db	185	EVI--SK--LYAGPEVDVMSQGVIIYALLCGTLFPD-----DENINLFKKIKGGIYTL	235		
QY	278	PSQWRFTPEALRMFORLLALEPERRGPAKEVFRFLKHELTSELRRRPSHRARKPP	333		
Db	236	PS---HLSALARDLIPMLVDDPMKATITREIRERORFOI-----RLPRLVLAIVPPP	283		

RESULT 9*
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C:Accession: A56009
R:Muranaka, I.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A:Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharo-
tase of Saccharomyces cerevisiae.
A:Reference number: A56009; MUID:94217693; PMID:8154654
A:Accession: A56009
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-511 <MUR>
A:Cross-references: GB:D26602; NID:G496384; PIDN:BAA05649.1; PID:G496385
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific prote
F:17-271/Domain: protein kinase homology <KIN>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

```

Query Match      14.8%; Score 332; DB 1; Length 511;
Best Local Similarity 31.9%; Pred. No. 2.9e-08;
Matches 92; Conservative 56; Mismatches 106; Indels 34; Gaps 10;

Qy 51 KHYELVRELKGNYGKVDLVVYKGTGTFKALKFKWNKSKT- ---LKNFLREVSTITNLS 106
Db 17 RNYKLGKTLGIGFGVKVTAETHLTGHKVAVKILNRKIKNMWEEKVRREIKLR-LFM 75

Qy 107 SPRIIKVDFVVFETEDCYEPAQYAPAGDLFDIIPQVGLPEDTVKRCVQOGLGLDFMH 166
Db 76 HPHIRLYEVVEPDSIYV-VMEYVKSSELFYIVEKRLQDEARFFQOIGSVYCH 134

Qy 167 GROLVHRDIKPNVLLFDRECRVKLADGWTR- -RVGCRVKRVSGTIPYTAPEVCAQR 224
Db 135 RNMVVRHDLKPNLLDSK--WNVKIA>DGLSNIMRDGHFLKTCGSPNYAAPEVI-SGK 191

Qy 225 ADGLAVDTGVDDWAFGLVLCVITGNFPEWAAAGADAFFEYVWRQGRILPGLPSQWRRF 284
Db 192 ---LYAGPEVDVNSCGVILYALLCGLTPFD-----DENIPNLFKIKGGMISLPS---HL 240

Qy 285 TEPALRMFQELLALEPERRGPAKEVFRFLKHELTSELRRRPPSHRKP 332
Db 241 SAGARDLIPRLMLVDP-----MKRNTIPEIRMEHPWQAHLP 276

```


C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C/Accession: T34074

R/Nelson, J.

A/Submitted to the EMBL Data Library, November 1995

A/Description: The sequence of *C. elegans* cosmid C01C4.

A/Reference number: Z21472

A/Accession: T34074

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-356 <N>

A/Cross-references: EMBL:U41025; PIDN:AAA82347.1; CESP:C01C4.3

C/Genetics:

A/Gene: CESP:C01C4.3

A/Introns: 18/3; 51/3; 109/1; 154/3; 205/1; 259/2; 288/3

C/Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 19.1%; Score 427.5; DB 2; Length 356;

Best Local Similarity 29.9%; Pred. No. 1.3e-12;

Matches 98; Conservative 75; Mismatches 136; Indels 19; Gaps 5;

48 DVTKHYELVRELKGTGYGVKVDLVVYKGTGTQKALFKVNSKTKLNFLEVSITNSLS 107

16 ILPNVYKLTGLIGSGFKVIAEHVVTGHKVAIKLNRKTKNMMEKVKREIKILR-L 74

105 SSSPFIKVFVFTEDCYVFAQBYAAGDLFDIIPQVGLPDTVKRCVQQLGLALDF 164

75 FMHPHIIRQYEVETTSIYV-VMEYVKSGLFDYVEKGRLEQDEARNFFQIIISGVEY 133

165 MHGROLVHRDIKPNVLLFDRECRVKLADGEMTR--RVGCRVKEVSGTIPTADEVCOA 222

134 CHRNVMVHRDLKPENLIL-DSRC-NIK-ADFGLSNWMDDGHFLTKTSCSPNAAEVI-S 190

223 GRADGLAVDTGVDVWAFGLIFCVLTGTFPWEAASGADAFEEFVRWGRGLPGPSQW- 281

191 GK--LYAGPEVDVWSCGVILYALICGL-LPPD-----DENIPNFKKIKGGIYILPSHLS 242

282 -----RRTEPALRM--FO-----RLALE-PERRGPAKEV----- 309

243 SEARDLIPRLITVDVKRITIPETIRQHWFOFTHLPYLAVSPDTPVEQAKKINEIVQEV 302

310 --FRELKHELTSERLRP-----SHRARKPGDRPPAGACPLRLAPGLKTKTV 355

303 VMGEDRNQVLESNRNTQNDATVYVLLDNRPVPSG-----YLS--EFQBT- 350

356 LTESGSGSRPAPPVAVGVSPLFPVVPVPEPGLAQGP--PGRTGRADKSKGO 413

351 -TDSSN-----PWRTEAGASPVGHWPVHAHDYHGLGARSQ 386

414 V 414

387 V 387

RESULT 4

S27966

Probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human

N/Alternate names: protein p78

C/Species: Homo sapiens (man)

C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 04-Mar-2000

A/Accession: S27966

R/Maheshwari, K.K.; Som, S.; Parsa, I.

submitted to the EMBL Data Library, January 1992

A/Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induce

A/Reference number: S27966

A/Accession: S27966

A/Molecule type: mRNA

A/Residues: 1-713 <MAH>

A/Cross-references: EMBL:M80359; NID:gl89511; PIDN:AAA59991.1; PID:gl89512

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:54-307/Domain: protein kinase homology <KIN>

F:62-70/Region: protein kinase ATP-binding motif

Query Match 15.2%; Score 340.5; DB 2; Length 713;

Best Local Similarity 24.4%; Pred. No. 1.6e-08;

Matches 112; Conservative 79; Mismatches 139; Indels 129; Gaps 17;

37 QALTILT-----LAASDVTKH---YELRELKGTGYGVKVDLVVYKGTGTQKALK 82

26 QEVTSRTSRSGARCNSTASCADQPHIGNTRLLTKTGKGNFAVKVLAHILTGREVAIK 85

93 FVNKSK---TKLKNFLREVSTNSLSSPFIKVFVVDVVFETEDCYVFAQYAPAGDLFDI 139

86 IIDKTQNLNPTSLQKLFREVIRMKIL-NEPNIVKLFEVI-ETQKTLYLIMEVASGKGVFDY 143

140 IPPQVGLPDTVKRCVQQLGLALDFMGRGRLVHRDIKPNVLLFDRECRVKLADFGMTR 199

144 LVAGRMKEKEARSKFRQVSAVQYCHQKRVHRDLKAENLL-DAD-WNIKIADFGFSN 201

200 R--VGCVRKRVSGTIPTADEVCOAGRADGLAVDTGVDVWAFGLIFCVLTGTFPWEAAS 257

202 EFTVGGKLDTCGSPFPAAPLFOGKYDG----PEVDVWVSLGVILYTLVSGSLPFDGQN 257

Query Match 15.4%; Score 344; DB 2; Length 512;

Best Local Similarity 28.5%; Pred. No. 8.5e-09;

Matches 120; Conservative 65; Mismatches 132; Indels 104; Gaps 20;

RESULT 3

T52633

serine/threonine-specific protein kinase (EC 2.7.1.-) AKIN11 [validated] - Arabidopsis

N/Alternate names: SNF1 protein kinase omolog AKIN11

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000

A/Accession: T52633

R/Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machid

Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999

A/Title: Regulatory interaction of pRL1 WD protein with Arabidopsis SNF1-like protein ki

A/Reference number: Z55116; MUID:99238528; PMID:10220464

A/Accession: T52633

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-512 <BHA>

A/Cross-references: EMBL:X99279; PIDN:CAA67671.1

A/Experimental source: cultivar Columbia

C/Genetics:

A/Gene: AKIN11

C/Function:

A/Description: EC 2.7.1.-; serine/threonine-specific protein kinase AKIN11 [validated, M

complements SNF1 mutations in yeast

C/Superfamily: AMP-activated protein kinase; protein kinase homology

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:54:36 ; Search time 20 Seconds
(without alignments)
2038.776 Million cell updates/sec

Title: US-10-016-985-2
 Perfect score: 2239
 Sequence: 1 MSVGCPEPEPRSLTCCGPG.....GRADSKGOVLATAIEICV 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 2833308

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	642.5	28.7	358	1	S71887		serine/threonine-s
2	427.5	19.1	356	2	T34074		hypothetical prote
3	344	15.4	512	2	T52633		serine/threonine-s
4	340.5	15.2	713	2	S27966		probable serine/th
5	336.5	15.0	512	1	JC1446		serine/threonine-s
6	336	15.0	504	2	T10449		probable serine/th
7	332.5	14.9	504	2	T07415		probable serine/th
8	332	14.8	510	2	T04145		serine/threonine p
9	332	14.8	511	1	A56009		serine/threonine p
10	332	14.8	513	1	S60303		serine/threonine-s
11	330.5	14.8	774	2	I48609		serine/threonine-s
12	328	14.6	460	2	S58882		probable serine/th
13	325	14.5	1032	2	D83637		protein kinase Cds
14	323	14.4	513	1	S60304		serine/threonine p
15	322.5	14.4	745	2	G01025		serine/threonine-s
16	318.5	14.2	473	1	S59941		serine/threonine p
17	318.5	14.2	1245	2	D86260		serine/threonine-s
18	317.5	14.2	576	2	T41587		protein T12C24.22
19	317.5	14.2	602	2	S72513		probable carbon ca
20	317.5	14.2	1398	2	T13741		FOG2 protein - yea
21	316.5	14.1	1192	2	T18611		hypothetical prote
22	316.5	14.1	1246	2	G89287		probable serine/th
23	315	14.1	520	2	G86414		protein H39E23.1 [
24	315	14.1	552	1	A53621		probable protein k
25	315	14.1	560	2	S51600		[hydroxymethylglut
26	314.5	14.0	562	2	T29558		phosphorylase kina
27	312	13.9	534	2	G89924		hypothetical prote
28	311.5	13.9	1051	1	JW0051		protein R02C2.1 [i
29	311	13.9	502	1	A41361		serine/threonine-s

ALIGNMENTS

RESULT 1

serine/threonine-specific kinase (EC 2.7.1.-), pk9.7 gastrula-specific - African clawed
S71887
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S71887
R;Shape, A.M.; Smith, J.C.
EMBO J. 15, 4556-4565, 1996
A;Title: Regulation of embryonic cell division by a Xenopus gastrula-specific protein kinase
A;Reference number: S71887; MUID:97042347; PMID:8887547

30	309.5	13.8	735	2	A53300	ribosomal protein
31	309	13.8	552	1	S51025	[hydroxymethyl]glut
32	310.5	13.8	552	1	S51025	serine/threonine-s
33	309	13.8	633	1	A26030	gik protein - chic
34	309	13.8	798	2	JC7500	protein kinase - m
35	307	13.7	481	2	I49072	probable serine/th
36	306	13.7	512	2	T07788	hypothetical prote
37	306	13.7	336	2	T21652	ribosomal protein
38	303	13.6	735	2	S15901	probable protein k
39	302	13.5	1081	2	S51899	GIN4 protein - yea
40	302	13.5	1142	2	S59359	hypothetical prote
41	301.5	13.5	1558	2	T29253	probable serine/th
42	300.5	13.4	440	2	T14736	ribosomal protein
43	299.5	13.4	740	2	I38556	probable Ca-calmod
44	298	13.3	580	2	T40939	serine/threonine k
45	295	13.2	856	2	I34631	probable protein k
46	295	13.2	435	2	E84707	

[illegible]

RESULT 2
T34074
hypothetical protein C01C4.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*